Remarks/Arguments

Claims 1-15 and 17 are pending in the application. Claims 1-9 and 15 have been withdrawn from consideration pursuant to a lack of unity objection. Claims 10-14 and 17 are therefore under consideration. Reconsideration is requested in view of the above changes and the following remarks.

Claims 10, 11 and 14 have been amended to recite that the pathogenic bacteria are extracellular pathogenic bacteria. Support for the amendment is found in the specification at page 6, lines 1-2, line 14, and lines 25-28. Claim 11 has been amended to recited that "wherein the antigenic peptide fragment which is also derived from the extracellular pathogenic bacteria" to make the scope of claim 11, in as far as it relates to the antigenic peptide fragment, consistent with that feature, as recited in claim 10.

Confirmation of Telephonic Interview

Applicant thanks Examiner for the courtesy extended to the undersigned counsel during the telephonic interview of Dec. 30, 2008. The Interview Summary is a complete record of the interview, supplemented as follows. No specific claim amendments were discussed. Applicant's counsel pointed to the fact that the grounds of rejection in the December 12, 2008 office action appeared to be repeated from the May 5, 2008 office action, notwithstanding applicant's arguments presented in the September 5, 2008 response to the May 5 office action.

Response to Rejection Under 35 U.S.C. § 102

Srivastava (US Patent No 5,961,979) – Rejection Under 35 U.S.C. § 102(e)

The Examiner has maintained the rejection of claims 10-14 and 17 as allegedly being anticipated by Srivastava (US Patent No 5,961,979). Applicant respectfully submits that the claims are not anticipated by Srivastava for the following reasons.

Notwithstanding, and solely in an effort to expedite prosecution of the present application, Applicant has amended the present claims 10-14 and 17 to introduce the feature that the pathogenic bacteria is an *extracellular* bacteria. Srivastava is not concerned with extracellular bacteria. Rather, Srivastava primarily relates to the production of heat shock

protein-antigenic peptide fragment complexes derived from mammalian eukaryotic cells which have been infected with intracellular bacteria. This is made clear in the teachings of Srivastava at column 1, lines 12-14 which read "*The application relates generally to the field of vaccine development, in particular, prophylactic and therapeutic vaccines effective against intracellular pathogens*". Further, column 4, lines 23-27 read "*It is an object of the instant invention to provide a safe subunit vaccine comprising a stress protein-peptide complex for administration to a mammal that is capable of inducing, by means of a cytotoxic T cell response, resistance to infection by a preselected intracellular pathogen*". In addition to the above referenced sections of Srivastava, there are numerous further references to the provision of mediating resistance against an intracellular pathogen throughout the description, for example at column 4, column 5, column 6 and column 7. In stark contrast, the teachings of Srivastava is silent with respect to extracellular pathogens. Accordingly, applicant respectfully submits that the feature of claims 10-14 and 17 which defines that the pathogen is an extracellular pathogen distinguishes these claims from Srivastava. Claims 10-14 and 17, as amended, are therefore not anticipated by the teachings of Srivastava.

In the office action, it is the Applicant's understanding that Examiner is implying that although Srivastava is directed to stress protein-peptide complexes which are derived from a mammalian cell, where the stress protein component is a mammalian heat shock protein, that the methods employed by Srivastava could *also* result in Srivastava providing a stress-protein complex wherein the stress protein component of the complex is derived from an intracellular bacteria. Accordingly, this would result in Srivastava providing a stress protein-peptide complex where both the stress protein and peptide component of that complex are derived from a bacterial cell.

Applicant submits, that if Srivastava were, in fact, to provide in amongst the (mammalian) stress protein / (intracellular bacteria derived) antigenic peptide fragment complexes, the production of which are the main intention of the teaching of Srivastava , further complexes which comprised a stress protein component and an antigenic fragment component which were *both* derived from the intracellular bacteria which was infecting the mammalian cell,

then such complexes *would not* be analogous to the induced stress protein-peptide complexes of the instant application which are derived from extracellular bacteria.

Specifically, the provision of a stress protein-peptide complex which is derived from extracellular bacteria is not contemplated by Srivastava, and further, Srivastava teaches of no method for such a complex to be obtained.

Furthermore, Applicant respectfully submits that the argumentation provided by Examiner set forth in the office action of December 12, 2008 as to how the teachings of Srivastava can be interpreted to disclose a bacterial derived stress protein complexed to a bacterial derived antigen peptide are inconsistent and therefore flawed.

Specifically, Examiner submits that Srivastava discloses the isolation and use of bacterial heat shock protein-bacterial peptide complexes for the eliciting of immune responses to those bacteria, through disruption of bacterial cells, and that this disclosure meets the requirements of Applicant's claims 10-14 and 17.

Claim 10, as amended, refers to heat shock complexes wherein "one or more endogenous complexes produced in-situ and extracted from the extracellular pathogenic bacteria between an induced heat shock protein which is derived from the extracellular pathogenic bacteria and an antigenic peptide fragment which is also derived from the extracellular pathogenic bacteria".

Under point 2 on page 4 of the present office action, Examiner submits that "*It is the position of the examiner that Srivastava discloses compositions that comprise Bacteria/E. coli heat shock protein-peptide complexes*" and "*The complexes comprises a heat shock protein and peptides that originate from the pathogen itself. (see col. 6, lines 12-13; col.6, lines 65-68 and col. 7, line 7).*"

The passages referred to by the Examiner Srivastava at col. 6, lines 12-13; col.6, lines 65-68 and col. 7, line 7 describe the potential for the *peptide* to be originated from the bacterial pathogen, but *do not* describe the use of a bacterial heat shock protein or stress protein. Applicant therefore respectfully submits that it is unclear as to where the Examiner, having read these passages, is erroneously drawing the conclusion that these passages describe the use of a bacterial heat shock protein-bacterial peptide complex. If Examiner is asserting that these passages teach of a complex where the heat shock protein is derived from a bacteria, then

Applicant respectfully submits that the bacteria is not an extracellular bacteria as recited in the claims of the instant application.

Furthermore, Examiner asserts later in point 2 that "*The bacterial heat shock protein (see Table 1 and definitions) is complexed together with an antigenic peptide fragment from a bacteria (see col. 7, line 7, "Chlamydia"), fungus, or protozoa, wherein the heat shock protein complex is isolated from natural sources (see col. 21, line 28)*". Applicant respectfully submits that in making this statement, and therefore asserting that Srivastava discloses stress protein-peptide complexes derived from a bacteria, that the Examiner has selected phrases from individual sentences, and assigned meanings to those phrases which are not supported by the meaning of the sentences as a whole. Hence, Examiner has incorrectly selected portions of the Srivastava description which are not inter-related and conjoined them to provide an argument and come to a conclusion that a stress protein-peptide complex derived entirely from a bacteria is taught. This conclusion is not justified from the totality of the reference's teaching, which must be considered as a whole. *See, W.L. Gore & Associates, Inc. v. Garlock, Inc.*, 220 USPQ 303, 311 (Fed.Cir.1983), *cert. denied*, 469 US 851 (1984).

For example, applicant respectfully draws Examiner's attention to the fact that the reference to Chlamydia cited above pertains to the use of an antigenic peptide *only* from this species, as evidenced by the preceding sentence from which the word Chlamydia is taken, and which reads "*Vaccines may be prepared that stimulate cytotoxic T cell responses against cells infected with intracellular parasites including, but not limited to, chlamydia and rickettsia*".

Applicant highlights the importance of the repeated references in Srivastava, such as that cited above, to the generation of a cytotoxic T cell (CTL) immune response by the complexes of Srivastava. A cytotoxic T cell response results from cytosolic antigen processing and subsequent presentation of antigen by major histocompatability complex class I (MHC I). This antigen processing pathway is therefore used to present foreign antigens which are present *within* cells to the immune system. Presentation of antigens via MHC I results in a cytotoxic T cell mediated immune response which destroys cells (in which the pathogens from which the antigens are derived). Such an immune response would be entirely inappropriate for preventing infection by an extracellular bacterial pathogen, as an extracellular bacterial pathogen would not be present

within a cell. In such instances, an antibody based "humoral" immune response must be mediated. This requires the extracellular pathogen to be taken up by a cell, typically a specialized antigen presenting cell, by a process such as endocytosis. Antigens present within the endosome are processed and presented by the antigen presenting cells by major histocompatability complex class II (MHC II) molecules. It should be noted that during the processing of extracellular antigens using the MHC class II pathway, the extracellular antigens are kept in vacuoles (first called endosomes, later called lysosomes) and therefore are separated from the cytosol. There is therefore no means for an antigen derived from an extracellular bacteria to pass into the cytosol and associate with a mammalian cell stress protein. Accordingly, the formation of a stress protein-peptide complex between a mammalian cell heat shock protein and an antigenic peptide fragment derived from an extracellular bacteria could not result.

This alternative antigen processing pathway does not result in an immune response being directed to the cells which present these foreign antigens (this would be pointless as the foreign antigens do not originate from within the antigen presenting cells), but rather results in an antibody-mediated immune response directed to the foreign antigens. Such a response includes a number of facets, such as neutralization, opsonization and complement activation, which target and destroy extracellular bacteria.

The disclosure of Srivastava refers only (and repeatedly) to the production of a cytotoxic T cell response. Srivastava is silent as to the use of the complexes derived therein for the production of an antibody-mediated humoral immune response. Applicant therefore submits that this provides further evidence that Srivastava did not intend to disclose, and does not disclose, stress-protein peptide complexes derived from extracellular bacteria.

Furthermore, in relation to the assertions made by Examiner, under point 2 of the present office action that "*The bacterial heat shock protein (see Table 1 and definitions) is complexed together with an antigenic peptide fragment from a bacteria (see col. 7, line 7, "Chlamydia"), fungus, or protozoa, wherein the heat shock protein complex is isolated from natural sources (see col. 21, line 28)*", in addition to the arguments set forth above, Applicant also respectfully asserts that the passages of Srivastava cited by Examiner relate to the *ex-vivo* combination of these naturally purified stress proteins with purified antigenic peptides, as evidenced by the

sentence as a whole, as found at column 21, lines 25-28 of Srivastava which reads "*As will be appreciated by those skilled in the art, the peptides either isolated by the aforementioned procedures or chemically synthesized may be reconstituted with a variety of naturally purified or recombinant stress proteins*". Consequently, the important inclusion of the word "reconstituted" clearly indicates that these complexes are formed in-vitro and not in-situ, as is required by claim 10, which refers to "*endogenous complexes produced in-situ and extracted from the pathogenic bacteria*".

Similarly, again under point 2, at lines 2-3 of the present office action, Examiner asserts "*The bacteria complexes are obtained through disruption of the bacterial pathogen*". Under point 4 on page 5 of the office action the Examiner asserts that "*...the heat shock protein/peptide complexes are disclosed in Srivastava to be obtainable from natural sources through disruption of the bacterial cells*". Examiner further asserts that "*Srivastava discloses compositions of heat shock proteins complexed with a peptide that originates from the bacterium in situ*". It is the position of the Applicant that Srivastava provides *no teaching* relating to the disruption of a pathogenic extracellular bacteria.

In her remarks, Examiner has not identified the passages of Srivastava from which the Examiner is drawing her conclusions. If, the Examiner is basing these conclusions on the cited passages referred to in the above paragraph (col. 7, line 7, "Chlamydia" and col. 21, line 28), the Applicant submits that the above two statements by the Examiner cannot describe or refer to a stress protein-peptide complex as recited in the amended claims of the instant application.

In view of the discussion presented above, Applicant submits it is clear that Srivastava does not specifically disclose stress protein-peptide complexes derived from bacteria. In particular, it is submitted that the sections of Srivastava cited by Examiner do not teach of a stress protein-peptide complex derived from bacteria. Without reference to specific teachings within Srivastava, Applicant can only speculate as to Examiner's reasoning for the above statements that "*The bacteria complexes are obtained through disruption of the bacterial pathogen*", "*...disclosed in Srivastava to be obtainable from natural sources through disruption of the bacterial cells*" and "*Srivastava discloses compositions of heat shock proteins complexed with a peptide that originates from the bacterium in situ*". However, it is the position of the

Applicant that the above statements by Examiner possibly suggest the Examiner's interpretation that the method of purification used by Srivastava to purify eukaryotic heat shock protein complexes from an infected eukaryotic cell results in the simultaneous co-purification of intracellular bacterial heat shock protein-bacterial peptide complexes, or that the two are indistinguishable.

However, according to the methods of Srivastava, this would *not* be the case. In particular, it is the opinion of the Applicant, that nowhere in Srivastava is the disruption of bacterial pathogens or bacterial cell lysis disclosed, or suggested. Specifically, the use of a dounce homogenizer, as taught at col. 13, line 63, would not be effective to cause the pathogenic bacteria which were infecting the host cell to be lysed or broken up. The subsequent ultracentrifugation step performed at 100,000 g, as taught at col. 13, lines 65-67 of Srivastava, would remove all bacteria from the system. It is well known in the art that such high centrifugation speeds is used to pellet viruses, bacteria and intracellular organelles, and therefore allow their separation from the supernatant which contains the stress protein-peptide complexes. Thus, the supernatant from this step that is loaded onto an affinity purification column, as taught at column 14, lines 1-2, would contain *only* eukaryotic proteins. Hence, neither bacterial heat shock proteins nor bacterial heat shock protein-peptide complexes would be present in the system following purification on the affinity column. Further, Applicant notes that Srivastava teaches the use of Con A SEPHAROSE (column 15, line 26). This purifies the glycosylated proteins. As bacteria lack Golgi apparatus, they do not possess the ability to glycosylate proteins. Accordingly, a stress protein-peptide complex derived from a bacteria would be aglycosylated (i.e. not glycosylated) and therefore would not be purified by the methods of Srivastava. As such, it is further submitted that Srivastava would not provide bacterial stress protein-peptide complexes, and in particular stress protein-peptide complexes derived from extracellular bacteria.

For these reasons, it is respectfully submitted that the claims do not lack novelty over the teachings of Srivastava.

Phipps et al. (1991, EMBO J., 10:1711-1722) – Rejection Under 35 U.S.C. § 102(b)

Examiner has maintained the rejection of claims 10-14 and 17 as allegedly anticipated by Phipps *et al.* ("Phipps"). Applicant respectfully submits that the claims are not anticipated by Phipps, for the following reasons.

Applicant submits that the teachings of Phipps do not disclose the subject matter of claims 10-14 and 17, as amended. In particular, Phipps does not disclose "*one or more endogenous complexes produced in-situ and extracted from the extracellular pathogenic bacteria between an induced heat shock protein which is derived from the extracellular pathogenic bacteria and an antigenic peptide fragment which is also derived from the extracellular pathogenic bacteria wherein production of the induced heat shock protein results from the exposure of the extracellular pathogenic bacteria to a stress-inducing heat shock stimulus*" as recited in claim 10.

Firstly, the disclosure of Phipps is concerned with a protein complex identified in the thermophillic archaebacteria *Pyrodictium occultum*. As taught by Phipps (page 1711, column 1) such bacteria are phylogenetically distinct from both the eubacteria and eukaryotes, and are amongst the most thermophilic organisms known to man, growing optimally at 105°C. These bacteria do not fall within the scope of the claims of the instant application, as amended, as bacteria described in Phipps are not pathogenic extracellular bacteria. They would not be able to survive in a human or animal host with a maximum body temperature of 40°C.

Secondly, and again referring to the features recited in claim 10, Phipps does not teach the claimed complexes ("*one or more endogenous complexes produced in-situ and extracted from the extracellular pathogenic bacteria between an* ***induced heat shock protein*** *which is derived from the extracellular pathogenic bacteria and an* ***antigenic peptide fragment*** *which is also derived from the extracellular pathogenic bacteria*"). In particular, at no stage in Phipps is a complex formed between an induced heat shock protein and an antigenic peptide fragment disclosed. The paper of Phipps attempts to characterize the identified protein complex, and on the basis of the identified structural and expression profile, Phipps theorizes that the complex is an ATPase complex where the experimental data leads Phipps to suggest "*The above observations suggest a possible relationship between Pyrodictium ATPase and the groEL family*

of proteins" (page 1720, column 1, lines 26-28). However, this is merely an assumption. Phipps states at page 1719, column 2, lines 53-54, that "*We do not yet know what the function of the ATPase complex is*". At no point in Phipps is the ATPase complex shown to bind to an antigenic peptide to form a stress protein-peptide complex, as is required by claim 10. Furthermore, in referring to a "complex" Phipps does not intend this to mean a complex of the ATPase protein and an antigenic peptide fragment.

Furthermore, applicant respectfully submits that the ATPase complexes which are the subject of the study by Phipps are not induced in response to heat shock, as required by the feature of claim 10 which reads "*induced heat shock protein which is derived from the extracellular pathogenic bacteria*". Rather, the disclosure of Phipps is concerned with a protein complex which is shown to be stable at high temperatures, and wherein the levels of this protein become "*enriched*" (see Abstract of Phipps), as opposed to being induced. This enrichment occurs because following heat shock, the levels of other proteins in the cytosol decrease and therefore the proportion of the identified ATPase complex increases. This is taught by Phipps at page 1717, column 1, paragraph 1 which reads "*It is clear that, under steady state conditions, the relative level of the ATPase increases with increasing growth temperature. Furthermore, the levels of most other proteins decrease, especially when the temperature is increased from 100°C to 108°C. At 108°C the ATPase constitutes 73% of the total soluble protein compared with 11% at 100°C and 6% at 90°C as judged by densitometry. Thus, at very high temperatures, the cytoplasmic protein consists predominantly of this single species*".

Phipps' Figure 9B shows the results of the protein composition of cells following heat shock. Phipps teaches at page 1717, column 2, 1st paragraph that "*4 h after shifting to the higher temperature, the ATPase accounts for a significantly larger fraction of the total soluble protein (lane d), ~35% as compared to 13% for the unshifted cells. We conclude that the ATPase complex is preferentially accumulated in response to heat shock in P. occultum, and that this leads to the establishment of a higher relative steady state level of the protein. These experiments do not allow us to decide whether this is due to an elevated rate of synthesis of ATPase, a reduced rate of synthesis of other cytoplasmic proteins or enhanced degradation of the latter*". The results shown in Figure 9B show no quantitative increase in the amount of the

ATPase complex in heat shocked cells over non heat shocked cells. Hence, Applicant submits that the level of ATPase has not actually increased in response to the heat shock, but rather the ATPase level has become more abundant within the cell, purely as a result of the levels of other proteins diminishing at the higher heat shock temperature.

Hence, there is no clear teaching in Phipps which indicates that the ATPase protein complex is, in fact, induced in response to heat shock, as required by the claims of the instant application. Furthermore, and as previously stated above, there is also no evidence in Phipps that the ATPase protein complex associates with an antigenic peptide fragment to form a stress protein-peptide complex.

To address the specific rejections set forth by Examiner in this regard, at point 6 of page 5 of the office action, Examiner states it is her position that Applicant's heat shock complex, as defined in the claims, results from the same process as that of Phipps, namely heat shock. In combining the observation that the proteins were produced in response to heat shock and further that the presence of 2 distinct bands, as present in certain lanes of Phipps Figures 9 and 10, the Examiner appears to have arrived at the conclusion that the complexes of Phipps anticipate the complexes recited in the claims of the instant application.

However, Examiner's attention is again drawn to the legend of Figure 9.A (page 1717, col. 1) which states "(A) *Steady state growth; SDS-PAGE of membrane-free lysates of P. occultum cells grown to exponential phase at different temperatures*" and to Page 1717, col.1, first paragraph, lines 4-7 "*we first examined the steady state levels of soluble proteins as a function of growth temperature by growing P. occultum cells to late exponential phase at 90°C, 100°C and 108°C*". Therefore, the data presented in Fig. 9.A do not represent heat shock complexes *induced* by heat shock, since the temperature was kept constant during the course of these experiments.

The fact that the protein complexes of Phipps are constitutively expressed and merely enriched following heat shock, as opposed to being *induced* following heat shock is a point of significance, as it is well known in the art that upon heat shock, heat shock proteins complex to a range of proteins/peptides that they, for the most part, do not complex to under steady state conditions. In this regard, the Applicant has shown in the examples provided in the instant

application (see example 3, page 15, lines 18-21 which states that "*The CZE profiles of the peptides eluted from the constitutive and heat-induced M. Tuberculosis HSPs were significantly different indicating that they carried distinct families of associated proteins*". In further support of this assertion, the Examiner's attention is drawn to Mogk *et al.*, 1999 (EMBO, vol. 18, No.24, pp 6934-6949, a copy of which is submitted herewith as Exhibit 1). Table III and Fig. 4 of Mogk *et al.*, 1999 definitively show the temperature-dependent association of protein substrates of the *E. coli* heat shock protein DnaK.

Therefore, a *constitutive* protein complex expressed under the steady state conditions as provided in Phipps cannot be considered by the Examiner to be the same as an *induced* heat shock protein provided under heat shock conditions, and as required by the claims of the instant application.

In Figure 9.B of Phipps, a heat shock stimulus was applied to the growing cells. However, upon analysis of the gel shown in Fig 9.B, Applicant can identify no quantitative difference between the level of protein shown in the complexes present in lanes b, c, and d. The only notable difference that can be seen in the bands of Figure 9.B is that in lane d, the other proteins previously present in lanes b and c are no longer present following heat shock. Therefore, it is submitted that the heat shock complexes in Fig 9.B have been *enriched* following heat shock, but not induced. This distinction is indeed made by the authors in the abstract of Phipps where it is stated that "*The basal level of the ATPase complex in the cell is high, and it becomes highly enriched as result of heat shock*" (Page 1711, abstract paragraph, lines 9-13).

Proteins that exhibit a high basal level are termed "constitutive" proteins, and such language is used by the authors of Phipps "*To investigate constitutive protein production as a function of growth temperature*" (Page 1721, col. 1, paragraph headed "Heat shock", lines 1-2). Further, the existence of such constitutive heat shock proteins that are not inducible by heat shock is well known in the art. See Arya *et al*, 2007 (Arya. R. *et al.* J. Biosci. 32(3), April 2007, p595-610 – a copy of which is submitted herewith as Exhibit 2) which at page 595, column 2, last line states that "*Many members of these Hsp families are present constitutively (heat shock cognates) in cells while some are expressed only after stress*". For an example specific to bacteria Examiner is referred to Seaton and Vickery (Seaton, B.L. and Vickery. L.E. Proc Nat.

Acad. Sci USA, Vol 91. p2066-2070, March 1994 – a copy of which is furnished herewith as Exhibit 3), which discusses the *E. coli* constitutive heat shock protein Hsc66 (Abstract- top of page 2066, col. 1, lines 9-11) "*In contrast to dnaK, however, the hsc gene lacks a consensus heat shock promoter sequence, and expression is not induced by elevated temp.*"

Therefore, although heat shock was applied to the complexes represented in Fig 9.B of Phipps, the Applicant respectfully submits that it is incorrect of Examiner to conclude that these complexes were *induced* in response to the heat shock applied. Indeed. Applicant refers again to the conclusion of Phipps, as fully recited above, which states that "*These experiments do not allow us to decide whether this is due to an elevated rate of synthesis of the ATPase, a reduced rate of synthesis of other cytoplasmic proteins or enhanced degradation of the latter*" (Page 1717, col.2, first paragraph, lines 11-14).

Under point 7 of page 6 of the office action, the Examiner refers to Fig. 10, lane h which shows proteins present in the heat shock complex of *E. coli*. By extension of the points raised above, Fig. 10 cannot show induced heat shock complexes either, for at least the reasons noted above, since the antibody in Fig. 10 is raised against the same non-induced/enriched heat shock complex of Fig. 9.A, and therefore could not be used to specifically bind to, and therefore distinguish, an induced ATPase complex, as opposed to an ATPase complex which is constitutively expressed. Phipps states: "*Membrane-free French press lysates of several archaebacteria, the eubacteria E. coli and Thermotoga maritama and the yeast Saccharomyces cerevisiae were electrophoresed in SDS gels and probed by immunoblotting with antiserum raised against the P. occultum ATPase complex*" (Page 1717, right hand col., second paragraph, lines 1-5).

Moreover, the Examiner asserts under point 7 on page 6 of the office action that as Fig. 10 shows a heat shock complex consisting of two bands, and Phipps et al teaches that the ATPase protein complex exists in the cell as two polypeptides of similar molecular weights, this prior art discloses a heat shock protein complex that comprises two proteins/peptides. The referenced bands do not relate to the stress protein and antigenic peptide fragments of a stress protein-peptide fragment complex, but rather to the ATPase complexes. In this regard, Examiner's attention is drawn to the legend of Fig. 11, which states that "*The complex is*

composed of an equimolar combination of 56 kd and 59 kd polypeptides which we suggest to alternate in each ring, each polypeptide corresponding to a single subunit". Therefore, the enriched ATPase has a double band banding pattern not because it is bound to a bacterial peptide, but because it consists of two separate subunits. Further, it is obvious from a reading of the Phipps document as a whole that the author is referring to the two sub-units mentioned in the legend to Fig. 11 when he states "*The reaction of the pair of bands in P. occultum (lane b) confirms that the protein exists in the cell as two polypeptides of similar mol. wt*" (Page 1717, right hand col., lines 7-10) as quoted by the Examiner on page 6 of the office action. Therefore, the data represented in Phipps Fig. 10, lane h shows a *single* heat shock protein, *enriched* following heat shock, which consists of *two subunits*. The figure therefore cannot disclose the induced heat shock protein-bacterial peptide complex of Applicant's claim 10.

Further, under points 7 and 8, Examiner suggests that the protein shown in Fig 10, lane h represents the *E. coli* heat shock protein, GroEL, and as there are two bands Fig. 10, lane h shows GroEL bound to a second protein/peptide, thereby disclosing the complexes of Applicant's claim 10. Notwithstanding the above points which demonstrate that Phipps merely discloses an enriched ATPase protein, as opposed to an induced heat shock protein, and also that the second protein/peptide represents the second subunit of the ATPase referred to in Fig. 11 and not a bacterial peptide complex, Examiner's attention is drawn to Phipps page 1718, col. 2, top paragraph, lines 9-10 which read "*Thus it is possible that one of the two cross-reacting bands corresponds to GroEL*". Therefore, a doubt exists on behalf of the author as to whether Fig. 10 Lane h *could* represent GroEL which doubt the Examiner has not taken into account. Moreover, the author then explains that in order to test this hypothesis, purified *Pyrodictium* ATPase (to which the antibody used in Fig. 10, lane h was raised) was reacted with an antibody raised against GroEL, which resulted in "*a very weak reaction in one of the two ATPase bands (data not shown)*" Page 1718, col. 2, top paragraph, line14. The low level of reactivity is likely due to the high conservation of structural epitopes amongst heat shock proteins, a matter conceded by the author in the last two lines of this paragraph "*Whether this reaction results from the presence of structurally similar epitopes in the two proteins is not clear*". It is obvious from this statement that Phipps does not consider the two proteins to be the same. It is respectfully submitted by the

Applicant that the Examiner cannot unambiguously conclude from these data and statements that the protein shown in Phipps Fig. 10, lane h definitively represents GroEL complexed to another peptide.

For these reasons, it is therefore respectfully submitted that the teachings of Phipps do not anticipate the claims of the instant application.

Wawrzynow *et al.* (1991, EMBO J., 9:1867-1877) – Rejection Under 35 U.S.C. § 102(b)

Examiner has maintained the rejection of claims 10 and 11 as allegedly anticipated by Wawrzynow *et al.* ("Wawrzynow"). Applicant respectfully submits that claims 10 and 11 are not anticipated by Wawrzynow, for the following reasons.

Under points 10 and 11 on page 9 of the office action, Examiner asserts that "*the disclosure of the complex of Wawrzynow et al meets all of the structural requirements of the claims*" on the ground that Wawrzynow discloses (i) an antibody raised to ClpX, as used in an ELISA assay, (ii) that ClpX is defined as a heat shock protein that is under heat shock regulation, and (iii) ClpX requires ATP for the formation of a complex with a peptide.

It is respectfully submitted by the Applicant that the teachings of Wawrzynow do not anticipate the claims of the present invention for the following reasons.

Firstly, Examiner's attention is drawn to the method used for the production of the ClpX heat shock protein in Wawrzynow, as taught at page 1875, last paragraph lines 2-3 "*The ClpX and ClpP proteins were purified essentially as described by Wojtkowiak et al (1993)*". Examiner's attention is specifically drawn to the Wojtkowiak *et al* (1993) document, a copy of which is furnished herewith as Exhibit 4 (Wojtkowiak D. *et al.* J. Biol Chem, Oct 25;268(30):22609-17) The Wojtkowiak *et al* document states at page 22610, column 1, last paragraph entitled "Purification of ClpX" -"*E. coli B178 cells were grown in LB medium (15 litres) with K_2HPO_4.....and sodium citrate(10 g) at 37°C*". Therefore, ClpX was not produced under heat shock in the experiments of Wawrzynow, but rather at *steady state* conditions of 37°C. Therefore, the complexes of Wawrzynow are not induced heat shock complexes within the meaning of Applicant's heat shock complexes of Claims 10-11.

Furthermore, Examiner's attention is again drawn to the reference made above to the teachings of Mogk *et al.*, 1999 (Mogk A. *et al.* EMBO, vol. 18, No.24, pp 6934-6949, Exhibit 1 hereto) where it is shown that upon heat shock, bacterial heat shock proteins complex to a range of proteins/peptides that they, for the most part, do not complex to under steady state conditions. Therefore, as no heat shock was applied in the production of the ClpX complexes disclosed in Wawrzynow, it is, in turn the case that the heat shock complexes of Wawrzynow are not "*complexes produced in-situ and extracted from the extracellular pathogenic bacteria between an induced heat shock protein which is derived from the extracellular pathogenic bacteria and an antigenic peptide fragment which is also derived from the extracellular pathogenic bacteria*" as recited in claims 10 and 11.

Secondly, Examiner's asserts that "*ClpX was shown to be immunoreactive with anti-ClpX antibodies in the ELISA immunoassay*" and that this disclosure in Wawrzynow meets all of the structural requirements of claims 10-11. In this regard Examiner's attention is drawn to teachings of Wawrzynow at page 1872, col. 2, paragraph 2, lines 9-11 which reads "*Following washing, the amount of ClpX bound to λO was detected by first incubating with rabbit anti-ClpX antibodies*". Therefore Wawrzynow discloses immuno-recognition of the ClpX heat shock protein only, and, importantly, *not* of the ClpX heat shock protein-λO peptide complex. Further, the λO peptide in this case is derived from a bacteriophage, not a bacteria, and so Wawrzynow discloses a bacterial heat shock protein-bacteriophage peptide complex, and not a "*an antigenic peptide fragment which is also derived from the extracellular pathogenic bacteria*", as required by claim 10 or claim 11.

Finally, the bacterial heat shock protein- bacteriophage peptide complex disclosed in Wawrzynow is produced under *in-vitro* conditions as evidenced on page 1872 col. 2, paragraph 2, lines 4-8, "*The λO protein was first allowed to fix in the wells of an ELISA plate, followed by blocking of the rest of the binding sites with excess BSA. The ClpX protein was then added.....*". Therefore, this complex was *not* produced *in-situ* as is specified claim 10. Moreover, the inability of Wawrzynow to isolate a ClpX- λO peptide produced *in-situ* is described at page 1872, column 2, final paragraph, lines 1-2 "*We were unsuccessful in our attempts to detect formation of a soluble λO-ClpX protein complex...*".

Therefore, for the above reasons it is respectfully submitted that there is no disclosure in Wawrzynow which anticipates claims 10 and 11

Motohashi *et al.* (Proc. Natl. Acad. Sci USA, Vol 96. pp7184-7189. June 1999) – Rejection Under 35 U.S.C. § 102(a)

Examiner has maintained the rejection against claim 10 as allegedly anticipated by Motohashi *et al.* ("Motohashi"). Applicant respectfully submits that claim 10 is not anticipated by Motohashi for the following reasons.

Under point 12 on page 10 of the office action the Examiner asserts that Motahashi discusses heat shock proteins in eubacteria, specifically T. thermophilus. In this regard the Examiner's attention is drawn to Page 7185, and the section beginning "Isolation of Proteins" in column one, the first line of which reads "*The TDnaK.J complex and TGrpE were expressed in E. coli...*"

Therefore, it is respectfully submitted by the Applicant that the bacterial heat shock protein-bacterial peptide complex of Motohashi, the TDnaK.J complex referred to above, does not meet the requirements of claim 10 for "*A composition for inducing an immune response to an extracellular pathogenic bacteria, the composition comprising one or more endogenous complexes produced in-situ and extracted from the extracellular pathogenic bacteria, which complexes are formed from an induced heat shock protein derived from the pathogenic bacteria*". This is because in the case of Motohashi, the pathogenic bacteria in which the complex is produced in-situ and extracted from is *E. coli*, yet the heat shock-peptide complex is, as the Examiner quite rightly states in the office action, derived from *T. thermophilus*.

Therefore the heat shock complexes of Motohashi are not produced *in-situ* and extracted from the *same* extracellular bacterial pathogen. For this reason, the Motohashi complexes cannot meet the requirements of Applicant's claim 10.

Applicant therefore respectfully submits that claim 10 is not anticipated by the teachings of Motohashi.

Langermann *et al.* (US Patent No. 6,500,434) – Rejection Under 35 U.S.C. § 102(e)

Examiner has introduced a new ground of rejection that claims 10-14 and 17 are allegedly anticipated by Langermann *et al.* (US Pat. 6,500,434) ("Langermann"). Applicant respectfully submits that the claims are not anticipated by Langermann for the following reasons.

Under point 15 on page 11 of the office action, Examiner asserts that "*Langermann et al claims FimC (Chaperone, heat shock protein)-FimH (peptide complexes) and a method of eliciting an immune response*. Examiner's attention is drawn to the fact that Langermann only discloses complexes composed of the periplasmic molecular chaperones FimC and PapD, which are not known to be involved in the heat shock response. These proteins are well known in the art to be involved in the transport of cytoplasmically synthesized proteins across the periplasm for extracellular pilus assembly, and to function differently from heat shock proteins. To support this assertion, applicant hereby refers Examiner to the teachings of Jones *et al.* 1993 (Jones C.H., *et al.* Proc. Natl. Acad. Sci USA (PNAS) Vol.90, Page 8397-8401, copy of which is submitted herewith as Exhibit 5) which teaches at page 8397, col. 1, lines 22-29 "*In contrast to cytoplasmic chaperones such as GroEL, DnaK, DnaJ, and SecB, which maintain their targets in highly unfolded conformations (10), PapD maintains target proteins in native-like conformations (8). In addition, periplasmic chaperones have an effector function, specifically targeting the subunits to outer membrane assembly sites for their incorporation into pili (11)*".

In this regard, it should be noted that there is a distinct and appreciable difference between being a chaperone and being a cytoprotectant. A chaperone protein is involved in bringing a protein to the targeted location for its specified purpose, whereas cytoprotection usually involves the modulation of particular protein activity or overall cell functionality, thereby leading to cell protection against injury or cell death. The skilled person would be fully aware of this distinction.

Further, additional differences between periplasmic chaperones and cytosolic heat shock proteins are noted by Miot and Betton 2004 (Miot and Betton, Microbial Cell Factories 2004, 3:4 (also available at URL: http://www.microbialcellfactories.com/content/3/1/4, a copy of which is provided herewith as Exhibit 6) which states at Page 4, column 1 that "*aside from specific chaperones, such as PapD involved in pilus assembly, few periplasmic chaperones have been*

identified, and there are no classical Hsp chaperones such as DNA or GroEL in this compartment. Indeed, as the periplasm lacks ATP, periplasmic chaperones must be mechanistically distinct from their cytoplasmic counterparts, most of which use ATP to drive their cycles of substrate binding and release".

Therefore, the claimed complexes of Langermann consist of non-heat shock induced periplasmic molecular chaperone complexes, whereby the formation of the complex is accomplished in *an ATP-independent* fashion. As such Applicant duly submits that the teachings of Langermann do not anticipate claims 10-14 and 17 as suggested by Examiner, in particular as Langermann does not teach the feature of claims 10 and 11 which require the stress protein-peptide complexes to be "*accomplished in an ATP-dependent reaction*".

Response to Obviousness-Type Double Patenting Rejection

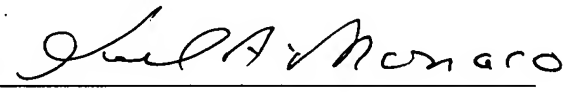
Applicant requests that the provisional non-statutory obviousness-type double patenting rejection be held in abeyance until claims have actually issued or are deemed allowing in copending Application No. 10/363,454.

Conclusion

The claims remaining in the application are believed to be in order for allowance. An early action toward that end is earnest solicited.

Respectfully submitted

CAMILO ANTHONY LEO SELWYN COLACO

BY 

DANIEL A. MONACO
Reg. No. 30,480
DRINKER, BIDDLE & REATH, LLP.
One Logan Square
18th and Cherry Streets
Philadelphia, PA 19103-6996
Tel.: (215) 988-3312
Fax.:(215) 988-2757
Attorney for the Applicant

EXHIBIT 1

The EMBO Journal Vol.18 No.24 pp.6934–6949, 1999

Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB

Axel Mogk, Toshifumi Tomoyasu, Pierre Goloubinoff[1], Stefan Rüdiger, Daniel Röder[2], Hanno Langen[2] and Bernd Bukau[3]

Institut für Biochemie und Molekularbiologie, Universität Freiburg, Hermann-Herder-Strasse 7, 79104 D-Freiburg, Germany, [1]Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel and [2]Hoffmann-La Roche AG, 4002 Basel, Switzerland

[3]Corresponding author
e-mail: bukau@uni-freiburg.de

We systematically analyzed the capability of the major cytosolic chaperones of *Escherichia coli* to cope with protein misfolding and aggregation during heat stress *in vivo* and in cell extracts. Under physiological heat stress conditions, only the DnaK system efficiently prevented the aggregation of thermolabile proteins, a surprisingly high number of 150–200 species, corresponding to 15–25% of detected proteins. Identification of thermolabile DnaK substrates by mass spectrometry revealed that they comprise 80% of the large (≥90 kDa) but only 18% of the small (≤30 kDa) cytosolic proteins and include essential proteins. The DnaK system in addition acts with ClpB to form a bi-chaperone system that quantitatively solubilizes aggregates of most of these proteins. Efficient solubilization also occurred in an *in vivo* order-of-addition experiment in which aggregates were formed prior to induction of synthesis of the bi-chaperone system. Our data indicate that large-sized proteins are most vulnerable to thermal unfolding and aggregation, and that the DnaK system has central, dual protective roles for these proteins by preventing their aggregation and, cooperatively with ClpB, mediating their disaggregation.

Keywords: chaperones/heat-shock response/Hsp70/protein denaturation/thermotolerance

Introduction

Misfolding and aggregation of proteins are major damaging consequences of stress situations such as heat shock and pathophysiological states (Morimoto *et al.*, 1994; Horwich and Weissman, 1997; Lindquist and Schirmer, 1999). The central cellular defense against such damage is molecular chaperones, which prevent aggregation, assist refolding and mediate degradation of misfolded proteins (Morimoto *et al.*, 1994; Hartl, 1996; Bukau, 1999). Chaperones can cooperate *in vitro* as part of a functional network in which 'holder' chaperones prevent aggregation of misfolded proteins, whereas 'folder' chaperones actively assist refolding (Langer *et al.*, 1992; Buchberger *et al.*, 1996; Freeman and Morimoto, 1996; Ehrnsperger *et al.*, 1997; Johnson and Craig, 1997; Veinger *et al.*, 1998). The contributions of individual chaperones to this folding network *in vivo* and the identity of the stress-sensitive cellular proteins remain unknown. Moreover, since cells have only a limited chaperone capacity to prevent protein aggregation under stress conditions (Craig and Gross, 1991; Tatsuta *et al.*, 1998; Tomoyasu *et al.*, 1998), it becomes important to determine the extent to which chaperones can resolubilize aggregates of proteins that escaped the protective function of holder chaperones.

For *Escherichia coli* it has been shown that the GroEL chaperone with its GroES co-chaperone assists folding of a subset of newly synthesized proteins and refolding of misfolded proteins (Hartl, 1996; Bukau and Horwich, 1998). Furthermore, overproduction of this chaperone system suppresses heat-induced aggregation of proteins in regulatory *rpoH* mutant cells that are defective in the synthesis of major chaperones (Gragerov *et al.*, 1992). The DnaK (Hsp70) chaperone and its DnaJ and GrpE co-chaperones participate in the folding of a subset of newly synthesized proteins (Deuerling *et al.*, 1999; Teter *et al.*, 1999), prevent protein aggregation at heat-shock temperatures (Gragerov *et al.*, 1992; Hesterkamp and Bukau, 1998) and refold misfolded proteins (Schröder *et al.*, 1993). The DnaK system has also been reported to dissolve aggregates of RNA polymerase and DnaA (Hwang *et al.*, 1990; Skowyra *et al.*, 1990; Ziemienowicz *et al.*, 1993), but it is inefficient in disaggregating other substrates including heat-aggregated firefly luciferase and malate dehydrogenase (Schröder *et al.*, 1993; Veinger *et al.*, 1998). Null mutations in the *clpB* gene encoding the *E.coli* Hsp104 homolog, ClpB, are retarded in the removal of a cell fraction assumed to be enriched in protein aggregates (Laskowska *et al.*, 1996a), but another study reported that Δ*clpB* mutants show no increase in protein aggregation at 46°C (Thomas and Baneyx, 1998).

The present work provides a systematic analysis of the roles of major chaperones of the *E.coli* cytosol in preventing and reverting heat-induced aggregation of proteins. This analysis was performed both *in vivo* and in cell extracts to allow independent validation of the *in vivo* results and to evaluate the usefulness of an *in vitro* approach. We identified for the first time a large set of thermolabile *E.coli* proteins, found to be enriched in large-sized proteins, and showed that the DnaK system is the single most effective chaperone in preventing their aggregation. Our findings thus provide the first identification of thermolabile *in vivo* substrates for an Hsp70 chaperone. Furthermore, we describe a powerful bi-chaperone system, consisting of ClpB and the DnaK system, which resolubilizes a wide variety of protein aggregates.

Results

The DnaK system is the most effective chaperone in preventing protein aggregation in extracts

The capacity of cytosolic chaperones to prevent heat-induced aggregation of *E.coli* proteins was determined in total soluble extracts of [^{35}S]methionine-labeled wild-type cells (MC4100). Heat treatment of extracts for 15 min at 45°C, a temperature within the growth temperature range of *E.coli*, caused denaturation and aggregation of 10–15% of the proteins despite the presence of endogenous chaperones (data not shown). We considered the possibility that many of the aggregating proteins are not thermolabile *per se*, but were co-precipitated through unspecific association with aggregating thermolabile proteins. We performed a seeding experiment in which [^{35}S]methionine-labeled cell extracts kept at 30°C were mixed with equimolar amounts of aggregating proteins of non-labeled cell extracts isolated at various time points after heat shock to 45°C. However, the mixing with aggregating proteins did not cause significant aggregation of the native, labeled proteins (Figure 1B), indicating that for most proteins, aggregation at 45°C is caused by thermal unfolding.

The aggregation of thermolabile proteins in cell extracts allowed us to test the potential of exogenously added chaperones in the presence of Mg^{2+}/ATP to prevent protein aggregation upon a subsequent heat shock. Chaperones were added to concentrations between 2.5 and 10 μM of monomers, a range corresponding to the concentration of aggregating proteins (~5–10 μM). Of the five chaperone systems tested (DnaK/DnaJ/GrpE, GroEL/GroES, ClpB, HtpG, IbpB), only the DnaK system suppressed protein aggregation in a concentration-dependent manner, with 90% efficiency reached at 7.5 μM DnaK (Figure 1A). For reasons explained below we tested a combination of the DnaK system and ClpB and found that it does not further increase the potential of the DnaK system to prevent protein aggregation (data not shown). We also investigated whether a crowded environment, as it exists in the *E.coli* cytosol (~170 mg proteins/ml *in vivo* versus 4 mg/ml in the extract), is sufficient to protect proteins from heat-induced aggregation. The presence of up to 20% (w/v) Ficoll 70 in cell extracts, which mimics the molecular crowding of the cellular environment (Martin and Hartl, 1997), caused no significant differences in protein aggregation (Figure 1A) and, in addition, did not alter the protection capacity of the DnaK or GroEL system (data not shown).

To judge whether the chosen concentration range of chaperones is physiological, we determined the chaperone levels in wild-type cells grown at 30°C or subjected to heat shock. The concentrations of all chaperones were in the low micromolar range and thus within the range used in the *in vitro* experiments described above. With respect to molar concentrations and the number of monomers, DnaK and GroEL were the most abundant chaperones at both temperatures (Table I), consistent with earlier determinations (Neidhardt and VanBogelen, 1987). However, since most chaperones are active as oligomers (hexamer for ClpB; dimer for HtpG; monomer for DnaK; tetradecamer for GroEL; unclear for IbpA/B), DnaK is at least 8-fold more abundant as active species compared

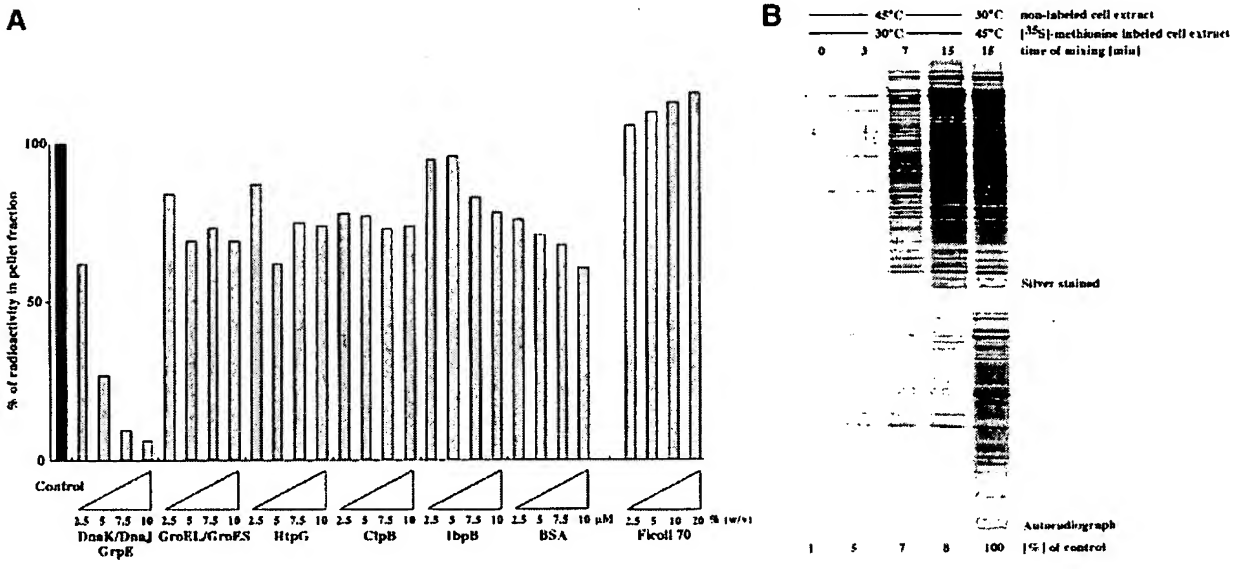


Fig. 1. (A) Prevention of heat-induced protein aggregation in *E.coli* cell extracts by chaperones. [^{35}S]methionine-labeled cell extracts were incubated for 5 min at 30°C and then 15 min at 45°C in the absence or presence of added chaperones, bovine serum albumin (BSA) or Ficoll 70 at the indicated concentrations [2.5–10 μM monomers or 2.5–20% (w/v)]. Aggregated proteins were isolated by centrifugation, boiled in 100 μl of SDS sample buffer and finally quantified by scintillation counting. Aggregation of heat-labile proteins in the absence of added chaperones was set at 100%. Protein ratios in the DnaK chaperone system were 1:0.2:0.1 (DnaK:DnaJ:GrpE) and 1:1 (GroEL:GroES) in the GroE chaperonin system. All experiments were carried out three times with <10% deviations. (B) Seeding of cell extracts kept at 30°C with aggregating proteins does not induce further protein aggregation. [^{35}S]methionine-labeled or non-labeled cell extracts were incubated at 30 or 45°C. At the time points indicated, extracts were mixed and incubated for a further 10 min at 30°C. Cell extracts kept at 45°C were cooled for 15 s on ice before addition. Insoluble cell fractions were isolated by centrifugation and analyzed by SDS–PAGE followed by silver staining. Dried gels were scanned using a phosphoimager (FLA-2000) and quantified by MacBAS software (Fuji film). The amount of labeled proteins in the pellet fraction of cell extracts kept at 45°C for 15 min was set at 100%.

with all other chaperones tested (Table I). Its high cellular concentration and high capacity to prevent protein aggregation in cell extracts qualifies DnaK as the central 'holder' chaperone of the *E.coli* cytosol.

The thermolabile proteins that showed increased aggregation in heat-treated cell extracts and were prevented from aggregation by the DnaK system were identified by two-dimensional (2D) gel electrophoresis of the aggregated protein fraction. This fraction consisted of ~250 protein species (Figure 2) corresponding to ~30% of all detected protein species. Addition of 5 μM DnaK and its co-chaperones (DnaJ, 1 μM; GrpE, 0.5 μM) to the extract before heat treatment prevented, at least partially, the aggregation of >90% of these protein species.

The DnaK system is the most effective chaperone system in preventing protein aggregation in vivo

The roles of chaperones in preventing thermal aggregation of proteins were determined *in vivo*. This analysis extends former studies that allowed only limited conclusions due to differences in the experimental and genetic conditions and partially controversial results (Gragerov *et al.*, 1992; Hesterkamp and Bukau, 1998; Thomas and Baneyx, 1998; Kedzierska *et al.*, 1999). We investigated the extent of heat-induced protein aggregation in wild-type and isogenic mutant cells carrying knockout mutations (*ΔclpB::kan*, *ΔhtpG::lacZ*, *ΔibpAB::kan*, *ΔdnaK52::cat*) or conditional mutations conferring temperature-sensitive growth at 42–43°C (*groEL44*, *groEL140*). The cells were grown at 30°C and then subjected to mild (60 min at 42°C) or severe (60 min at 45°C) heat-shock treatment. Within the time course of the experiment, none of the mutant strains showed growth defects as judged by increases in optical density of the cultures. *ΔdnaK52::cat* and *groEL* mutant cells stopped growing only 2–3 and 1–2 h after temperature upshift to 42 and 45°C, respectively. *ΔhtpG::lacZ* mutants had a reduced growth rate after heat shock to 45°C, as reported previously (Bardwell and Craig, 1988).

Heat-shock treatment at 42°C caused strong protein aggregation in *ΔdnaK52::cat* mutants (Figure 3), but no detectable aggregation in the wild-type and other mutant cells tested (data not shown). Treatment at 45°C caused even stronger protein aggregation in *ΔdnaK52::cat* mutants, minor accumulation of aggregated proteins in *groEL44*, *groEL140* and *ΔclpB::kan* mutant cells, and no detectable protein aggregation in *ΔhtpG::lacZ* and *ΔibpAB::kan* cells (see Figure 8; data not shown). In the case of the conditional *groEL* mutants we can not exclude the possibility that some remaining activity of GroEL makes a potential contribution of GroEL to aggregation prevention. We consider this contribution as minor, given that the onset of growth arrest of these mutant cells at 42°C was earlier than that of *ΔdnaK52::cat* mutant cells and the GroEL system was inefficient in aggregation prevention in cell extracts.

Our *in vivo* data are consistent with and confirm the results of the *in vitro* experiments described above. DnaK is thus the most effective chaperone to prevent aggregation

Table I. Cellular levels of cytosolic chaperones of *E.coli*

Protein	Concentration [μM] (monomers)		Number per cell (monomers/active species)	
	30°C	42°C	30°C	42°C
ClpB	8.7	19.4	3000/500	6700/1115
HtpG	5.8	21.2	2100/1050	7700/3850
DnaK	27	54	9900/9900	20 000/20 000
GroEL	47	99	17 200/1230	36 200/2590
IbpA/B	1.5	11.2	600/<600	5200/<5200

Wild-type cells (MC4100) were grown logarithmically in LB medium at 30°C and aliquots were further incubated at 42°C for 30 min. Aliquots of total cell extracts were subjected to SDS–PAGE followed by immunoblot analysis using chaperone-specific antisera. Serial dilutions of purified proteins served as a standard and allowed quantification of the signals in the linear range. The number of active species per cell was calculated on the basis of the known oligomeric state of the chaperones. IbpB purifies as a heterogeneous mixture of monomers and oligomers.

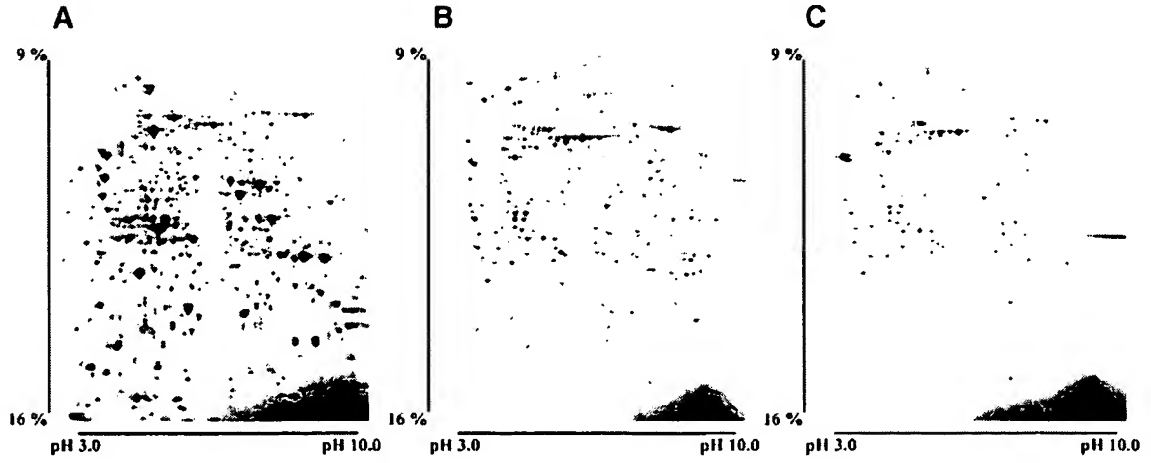


Fig. 2. Heat-induced protein aggregation in cell extracts: dependence on protein size and general protection by DnaK. Cell extracts were heated to 45°C for 15 min in the absence or presence of exogenously added DnaK/DnaJ/GrpE. Soluble and insoluble fractions were separated by centrifugation and analyzed by 2D gel electrophoresis. (**A**) Total soluble proteins of wild type (MC4100). (**B**) Aggregated proteins after heat treatment for 15 min at 45°C. (**C**) Protection of heat-labile *E.coli* proteins from aggregation by the DnaK chaperone system *in vitro*. DnaK (5 μM), DnaJ (1 μM), GrpE (0.5 μM) and ATP (10 mM) were added prior to heat shock. Aggregated heat-labile proteins were isolated and analyzed by 2D gel electrophoresis.

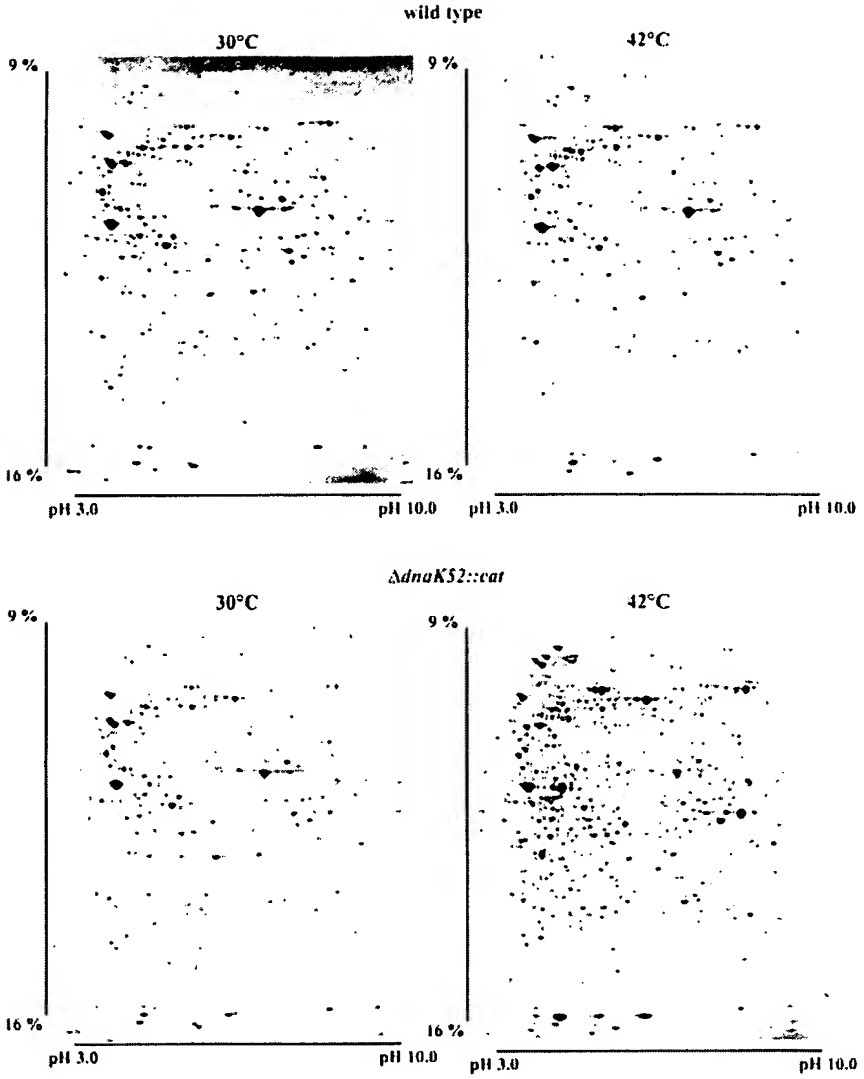


Fig. 3. Two-dimensional gel electrophoresis of insoluble cell fractions of *E.coli* MC4100 wild-type and Δ*dnaK52::cat* mutant cells, isolated at 30°C or 60 min after heat shock to 42°C. Cultures of wild-type and mutant strains were grown in M9 minimal medium supplemented with 19 L-amino acids except L-methionine at 30°C to logarithmic phase and shifted to 42°C for 60 min. Insoluble proteins were analyzed by 2D gel electrophoresis. Pellet fractions contain membrane proteins and aggregated protein material.

of thermolabile proteins in the cell, and this activity explains the temperature-sensitive growth phenotype of Δ*dnaK52::cat* mutant cells. The inability of the other chaperones to prevent protein aggregation efficiently *in vitro*, and the lack of strong protein aggregation phenotypes of the mutants lacking functions of these chaperones, indicate that they play only minor roles in aggregation prevention. This finding was further supported by tests for complementation of the temperature-sensitive growth phenotype of Δ*dnaK52::cat* mutant cells by overproduction of other chaperones. Isopropyl-β-D-thiogalactopyranoside (IPTG)-controlled overproduction of GroEL/GroES, ClpB, HtpG and IbpA/B did not restore the ability of Δ*dnaK52::cat* mutant cells to grow at 42°C (Table II). Partial complementation was observed at 40°C by overproduction of GroEL/GroES. This finding was consistent with the observation that a massive overproduction of GroEL/GroES (~30% of total protein) partially suppressed protein aggregation at 42°C in Δ*dnaK52::cat* cells, while massive overproduction of all other chaperones tested had only minor (IbpA/B) or no (HtpG, ClpB) effects in preventing heat-induced protein aggregation (data not shown). It is important to emphasize that these effects were only observed upon overproduction of the chaperones to very high levels and are therefore considered unphysiological.

Identification of natural DnaK substrates

We first analyzed the proteins that aggregated in Δ*dnaK52::cat* mutant cells after shift to 42°C by 2D gel electrophoresis of the insoluble fractions (Figure 3). These fractions contained membrane proteins and protein aggregates if present. The membrane proteins served as an internal loading control since their amount stayed constant for all *E.coli* strains and temperatures tested. Removal of most membrane proteins from the insoluble

Table II. Complementation of the temperature-sensitive growth phenotype of *ΔdnaK52::cat* mutants by plasmid-encoded chaperone alleles

Allele	Colony formation											
	30°C		40°C					42°C				
	0	500	0	50	100	250	500	0	50	100	250	500
dnaK, dnaJ	+	+	−	(+)	+	+	+	−	(+)	+	+	+
groEL, groES	+	+	−	−	(+)	(+)	−	−	−	−	−	−
clpB	+	+	−	−	−	−	−	−	−	−	−	−
hlpG	+	+	−	−	−	−	−	−	−	−	−	−
ibpA, ibpB	+	+	−	−	−	−	−	−	−	−	−	−

ΔdnaK52::cat mutants that carry pDMI,1 (*lacI*^{q}) as well as IPTG-inducible expression vectors (pUHE derivatives) (Lutz and Bujard, 1997) carrying the given chaperone alleles were tested for growth at 30, 40 and 42°C at IPTG concentrations as indicated in the top row (μM). Growth was assayed by determining the ability of the cells to form colonies on LB/Ap/Km agar plates. +, normal number and size of colonies; (+) reduced number and size of colonies; −, no colonies.

fractions in the presence of 2% (v/v) NP-40 during the extraction did not change the amount or pattern of the aggregated proteins (data not shown).

For wild-type cells kept at 30°C or incubated for 60 min at 42°C, the insoluble fractions showed no differences in the overall pattern of the proteins, of which the most prominent are membrane proteins (Figure 3). Only a slight increase in the amount of insoluble proteins was detected in *ΔdnaK52::cat* mutant cells kept at 30°C, consistent with earlier findings that DnaK is dispensable for folding of the majority of newly synthesized proteins at 30°C (Hesterkamp and Bukau, 1998). Increases in spot intensities compared with wild type were reproducibly detected for a few cytosolic proteins, including AceE (E1 component of pyruvate dehydrogenase), β and β′ subunits of RNA polymerase, PurL (formylglycinamidine synthase) and CarB (carbamoylphosphate synthetase). In heat-treated *ΔdnaK52::cat* mutants, ~10% of the amount of pre-existing cytosolic proteins of *ΔdnaK52::cat* mutant cells aggregated within 1 h of incubation of the cells at 42°C. Of the ~800 proteins resolved in total soluble cell lysate, the aggregating proteins comprised 150–200 species, defined as spots whose intensity is at least 4-fold increased compared with the 30°C *ΔdnaK52::cat* and wild-type controls. Thus, ~15–25% of the detected soluble protein species showed increased aggregation in *ΔdnaK52::cat* mutant cells at 42°C. The degree to which individual proteins aggregated varied between 30 and 70% of the particular spot.

The aggregated proteins were identified by subjecting 2D gel spots to Lys-C digestion followed by mass spectrometry of the peptides. In total lysates of *ΔdnaK52::cat* mutant and wild-type cells grown at 30°C, we identified 268 proteins of the soluble fraction and 14 proteins of the insoluble fraction that were membrane proteins. Fifty-seven of the identified soluble proteins specifically aggregated in *ΔdnaK52::cat* cells upon incubation at 42°C (Table III). These proteins were exclusively cytosolic and did not include any of the 28 periplasmic proteins or 14 membrane proteins identified.

We then determined the relationship between the aggregated proteins of heat-treated *ΔdnaK52::cat* mutant cells (Figure 3) and those of heat-treated cell extracts (Figure 2B) using comparative spot matching analysis and mass spectrometry. There was a good overall match between the two protein populations. More than 80% of the aggregating proteins identified *in vivo* and in cell extracts were identical (see Table III for summary); moreover, many of these proteins were identical to those co-immunoprecipitated with DnaK at 45°C (see below; Figure 4; Table III). These findings indicate that most of the proteins that aggregate in *ΔdnaK52::cat* mutant cells at 42°C are the major thermolabile proteins of *E.coli* as identified in cell extracts and are natural substrates for the DnaK chaperone system.

We defined the features of the aggregation-prone, thermolabile proteins that are substrates for DnaK. We compared the proteins that aggregated in *ΔdnaK52::cat* mutant cells at 42°C with total soluble proteins of cells kept at 30°C, and included the 2D gel spots identified by mass spectrometry and additional detectable spots evaluated by ImageMaster software (Pharmacia). No correlation was observed with respect to cellular function, pI value or oligomeric state. In contrast, a strong correlation existed with respect to molecular weight. The total soluble proteins detected in the 2D gels showed a broad size distribution with an average molecular weight of 42 kDa (Figure 5A). These characteristics were similar to those calculated on the basis of the entire *E.coli* genome (Netzer and Hartl, 1998), taking into account that very small proteins <10 kDa can not be resolved in the 2D gel system used. The aggregated proteins isolated from heat-treated *ΔdnaK52::cat* mutant cells, or from heat-treated cell extracts, showed a 2- to 4-fold increase in the relative number of large proteins (>70 kDa) and a 2- to 3-fold decrease in the relative number of small proteins (<30 kDa). Similar results were obtained when the intensities of spots were taken as a basis for this evaluation (data not shown). These characteristics have dramatic consequences with respect to the aggregation behaviour of large- and small-sized proteins. Approximately 80% of the soluble proteins >90 kDa that were detected in the soluble fraction of wild-type cells showed increased aggregation in heat-treated *ΔdnaK52::cat* mutants and extracts of wild-type cells (Figure 5A). In contrast, only 9 and 18% of the identified soluble proteins <20 and <30 kDa, respectively, showed a tendency to aggregate. Thus, the most striking finding of this analysis is that most large-sized cytosolic proteins are thermolabile even under mild heat-shock conditions, but are prevented from aggregating by the DnaK system.

Table III. Aggregation-prone *E.coli* proteins and their interaction with the DnaK system and ClpB

	Name	Size (kDa)	Function	Aggregation		CoIP		Disaggregation	
				in vivo	extract	*in vivo*	extract	*in vivo*	extract
1	RpoB	151	RNA polymerase	+	+	(+)	+	+	+
2	PutA	145	proline dehydrogenase	+	+			+	+
3	PurL	142	formylglycinamidine synthase	+	+		(+)	+	+
4	NarG	141	nitrate reductase (respiratory chain)	+	+			+	+
5	MetH	137	methionine biosynthesis	+	+	+	+	+	+
6	NifJ	130	oxidoreductase (respiratory chain)	+	+			+	+
7	TrcF	130	transcription repair coupling factor	+	+			+	+
8	CarB	118	carbamoylphosphate synthetase	+	+		(+)	+	+
9	SucA	106	E1 component of oxoglutarate dehydrogenase	+	+	+	+	+	+
10	AceE	100	E1 component of pyruvate dehydrogenase	+	+	+*	+*	+	+
11	AdhE	96	alcohol dehydrogenase	+	+	+	+	+	+
12	AcnB	94	aconitase	+	+	+	+	+	+
13	Lon	88	ATP-dependent protease	+	+	+		+	+
14	MetE	85	methionine biosynthesis	+	+	+	(+)	+	+
15	YghF	82	unknown	+	+	+		+	+
16	EF-G	78	elongation factor of translation	+	+	+	(+)	+	+
17	AckA	77	acetyltransferase	+	+	+*	+	+	+
18	ThrS	75	threonine tRNA synthetase	+				+	
19	Dxs	68	transketolase	+		+		+	
20	SfcA	66	probable malate oxidoreductase	+	+	+	+	+	+
21	TypA	66	unknown, EF-G homolog	+	+	+	+	+	+
22	AceF	66	E2 component of pyruvate dehydrogenase	+				+	
23	SdhA	65	succinate dehydrogenase	+	(+)			+	
24	FumA	61	fumarase	+	+	+		+	+
25	GuaA	59	glutamine amidotransferase	+	(+)			+	+
26	LysS	57	Lys-tRNA synthetase	+	+			+	+
27	AsnS	53	asparagine-tRNA synthetase	+	(+)			+	
28	GlnA	52	glutamine synthetase	+				+	
29	ImdH	52	IMP dehydrogenase	+	(+)			+	
30	HslU	50	subunit of ATP-dependent protease	+	+	+*	+	+	+
31	ClpX	47	subunit of ATP-dependent protease	+	+		+	+	+
32	ThrC	47	threonine biosynthesis	+	+	+*	+	+	+
33	DeoA	47	thymidine phosphorylase	+	+		+	+	+
34	Rho	47	termination of transcription	+	+			+	+
35	MurA	45	cell wall synthesis	+			(+)	+	
36	SerA	44	3-phosphoglycerate dehydrogenase	+	+	(+)	(+)	+	+
37	Tgt	43	quenine-tRNA ribosyl transferase	+	(+)	(+)	(+)	+	
38	Glf	43	LPS synthesis	+	+	+	+	+	
39	CarA	42	carbamoylphosphate synthetase	+	(+)			+	
40	RfbB	41	DTDP-glucose-4,6-dehydratase	+	+		(+)	+	+
41	Pgk	41	phosphoglycerate kinase	+				+	
42	FtsZ	40	cell division	+	+			+	+
43	YchF	40	probable GTP-binding protein	+	+		+	+	+
44	RpoA	37	RNA polymerase	+	+		(+)	+	+
45	AsnA	37	asparagine synthetase	+	+	(+)	(+)	+	+
46	GAP-DH	36	3-glyceraldehyde dehydrogenase	+	+		+	+	
47	MhpE	34	3-hydroxyphenylpropionate degradation	+	(+)	+	+	+	+
48	EntB	33	isochorismatase	+	+	+	+	+	+
49	MetF	33	methionine biosynthesis	+	(+)		(+)	+	
50	GatY	31	fructose bisphospate aldolase	+	+		+	+	
51	MinD	30	cell division	+	+			+	+
52	GpmA	29	phosphoglycerate mutase	+	+	(+)		+	+
53	ArcA	27	response regulator (oxygen control)	+	+	+		+	+
54	UbiB	26	flavoprotein oxidoreductase	+	+			+	
55	PyrH	26	UMP kinase	+	+	+		+	
56	NusG	21	antiterminator of transcription	+	+			+	
57	IbpB	16	small HSP (chaperone)	+				+	

Fifty-seven proteins were identified by mass spectrometry based on their specific aggregation in *ΔdnaK52::cat* mutant cells (BB1553, grown at 30°C to log phase) that were incubated at 42°C for 60 min (Aggregation, *in vivo*). This list comprises proteins identified in cells grown in LB or M9 minimal media, on the basis of at least four independent experiments each. Protein names, sizes and functions are indicated. These proteins were compared with the aggregated protein fraction of extracts of wild-type cells (MC4100) subjected to a 30–45°C heat shock for 15 min (Figure 2), and proteins co-immunoprecipitated with DnaK-specific antiserum from wild-type spheroplasts (MC4100) (CoIP, *in vivo*) and cell extracts (CoIP, extract) subjected to heat treatment (Figure 4). In addition, the disaggregation of heat-aggregated DnaK substrates by ClpB and the DnaK chaperone system in cell extracts (Disaggregation, extract) (Figure 7) and in *ΔdnaK52::cat* mutant cells (Disaggregation, *in vivo*) are indicated (Figure 9). (+), only minor amounts were detected. +*, proteins could also be co-immunoprecipitated at 30°C.

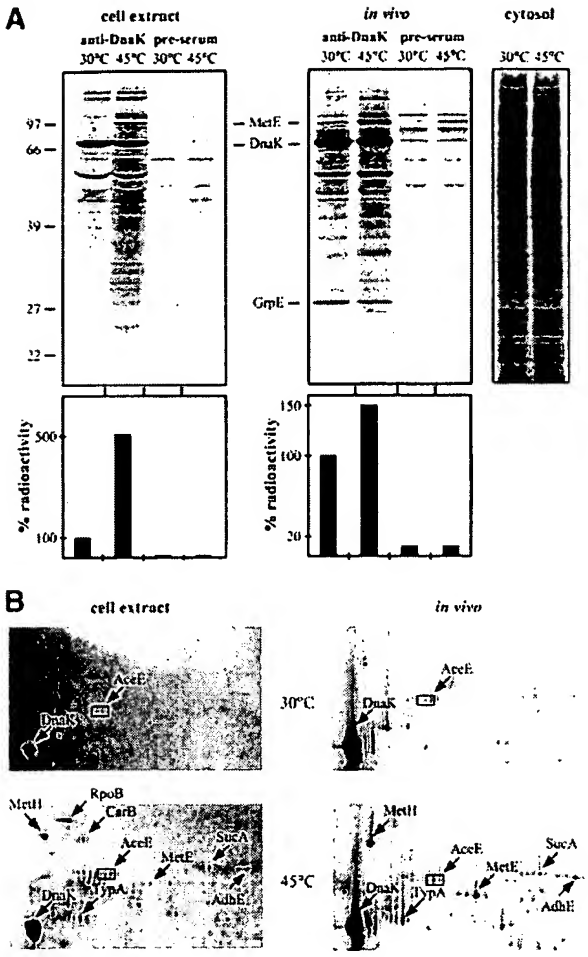


Fig. 4. Temperature-dependent association of protein substrates with DnaK. [^{35}S]methionine-labeled cell extracts (4 mg/ml, 10 mM ATP) were incubated at 30°C or shifted to 45°C for 15 min in the presence of exogenously added 7.5 μM DnaK, 1.5 μM DnaJ and 0.75 μM GrpE. [^{35}S]methionine-labeled spheroplasts (*in vivo*) were incubated at 30°C or shifted to 45°C for 5 min and lysed by addition of EDTA (100 mM) and sonification. All reactions were subjected to co-immunoprecipitation using polyclonal antibodies raised against the ATPase domain of DnaK or pre-serum as a control. Immunoprecipitated material was analyzed by one- and two-dimensional gel electrophoresis. (**A**) Autoradiograph of immunoprecipitated material, isolated by incubation with DnaK-specific antiserum or pre-serum. The lower panel shows quantification of the immunoprecipitated material by scintillation counting and MacBAS software. The amount of radiolabeled immunoprecipitated material at 30°C was set at 100%. [^{35}S]methionine-labeled cytosolic extracts from *E.coli* MC4100 cells are given. (**B**) Identical parts of 2D gels of immunoprecipitated material. Upper part, co-immunoprecipitation at 30°C; lower part, co-immunoprecipitation at 45°C. Identified proteins are indicated.

DnaK interacts with a broad variety of thermolabile proteins in extracts and in vivo

We investigated by co-immunoprecipitation whether the protective role of the DnaK system in cell extracts and *in vivo* relies on direct association with the aggregation-prone proteins. This approach is expected to detect only a subfraction of the DnaK substrates, given the highly transient nature of DnaK–substrate complexes and unavoidable losses during co-immunoprecipitation that limit the detectability of less abundant substrates. First, [^{35}S]methionine-labeled extracts supplemented with

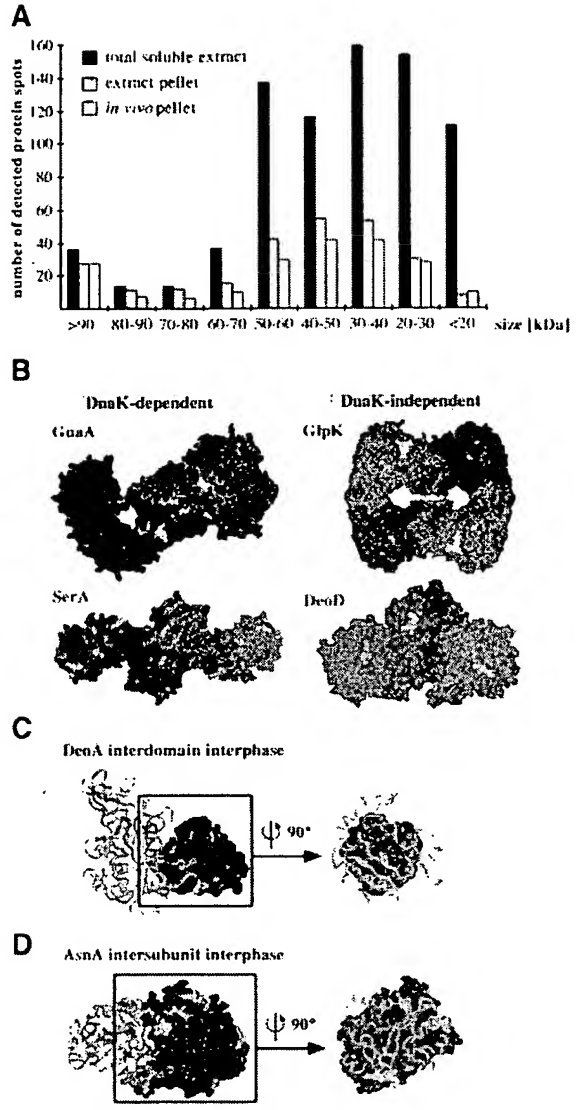


Fig. 5. (**A**) Protein aggregation in wild-type extracts and Δ*dnaK52::cat* mutant cells depends on protein size. Two-dimensional gels representing total soluble proteins of wild type (MC4100), aggregated proteins of wild-type cell extracts after incubation at 45°C and aggregated proteins in a Δ*dnaK52::cat* mutant after heat shock to 42°C were analyzed by ImageMaster software (Pharmacia Biotech). Molecular masses of protein spots were calculated following manufacturers' protocols using known masses of identified proteins. Total numbers of detected protein spots are compared within defined molecular weight groups. (B–D) Examples of thermolabile DnaK substrates with increased surface hydrophobicity. Potential DnaK-binding sites had been identified using the algorithm described previously (Rüdiger *et al.*, 1997b) and localized in the three-dimensional structures of proteins that are aggregation prone (GuaA, SerA, DeoA, AsnA) or stable (GlpK, DeoD) in the absence of DnaK at 42°C. The residues of the hydrophobic core of DnaK-binding sites with a prediction value of less than –4 are shown in gray, the hydrophobic side chains are shown in red. (**B**) GuaA and SerA expose several patches of hydrophobic side chains of potential DnaK-binding sites. GlpK and DeoD expose only a few sites. (**C**) The N-terminal domain of DeoA (blue space filling) exposes DnaK sites on the contact site to the C-terminal domain (green ribbon). (**D**) The homodimeric AsnA exposes DnaK-binding sites on the dimerization interphase (one subunit is shown as blue space fillings, the other as green ribbon).

Mg^{2+}/ATP were incubated at 30 or 45°C in the presence of added DnaK, DnaJ and GrpE. When incubated at 30°C, DnaK was co-immunoprecipitated with only few labeled proteins (Figure 4A). Seventy-five to eighty-five percent of the labeled protein corresponded to endogenous DnaK itself. After heat treatment, the amount of co-immunoprecipitated proteins increased by 5- to 6-fold compared with the 30°C control. Secondly, [^{35}S]methionine-labeled spheroplasts of *E.coli* wild-type cells, which are readily lysed and immediately accessible for co-immunoprecipitation, were subjected to heat treatment at 45°C. Before the temperature upshift an excess of unlabeled L-methionine was added to avoid detection of newly synthesized, nascent polypeptide chains, which may associate with DnaK. Seventy percent of the labeled protein corresponded to endogenous DnaK itself. After heat treatment, a 1.5-fold increase in the amount of co-immunoprecipitated proteins was observed compared with the 30°C control (Figure 4A). The lower number of co-immunoprecipitated proteins in heat-treated spheroplasts compared with heat-treated cell extracts may be due to differences in the DnaK concentrations. In the extracts, we used DnaK concentrations that were sufficient to allow complete prevention of protein aggregation. In contrast, the *in vivo* concentration of DnaK is limiting for aggregation prevention at 45°C, as shown by substantial transient aggregation of proteins in wild-type cells subjected to a 45°C heat treatment (see Figure 8).

Two-dimensional gel electrophoresis of the co-immunoprecipitated proteins revealed that they represent a broad spectrum of species (Figure 4B shows segments of the gels). Spot matching with 2D gels of heat-treated cell extracts and spheroplasts showed that many of the proteins protected by DnaK from thermal aggregation in extracts and *in vivo* were found associated with the chaperone after a 45°C treatment (Table III). We noted that AceE, which showed slightly enhanced aggregation in *ΔdnaK52::cat* mutant cells even at 30°C, could be co-immunoprecipitated at that temperature, pointing to a possible role of DnaK in *de novo* folding or assembly of the pyruvate–dehydrogenase complex. Together, these co-immunoprecipitation experiments detected >60 proteins associated with DnaK predominantly at high temperature both in spheroplasts and extracts, of which >25 were identified by comparative spot matching with reference gels. These proteins represent a subpopulation of those detected as aggregated proteins in *ΔdnaK52::cat* mutant cells and cell extracts (~200), of which 57 proteins were identified by mass spectrometry and no additional protein species were observed. The similarity of the patterns of identified DnaK substrates in spheroplasts and in extracts further strengthens the physiological relevance of the results obtained with extracts. In consideration of the mentioned difficulty in catching many of the unstable DnaK–substrate complexes, these findings indicate that the prevention of aggregation of thermolabile proteins by the DnaK system involves at least in many cases a direct interaction of DnaK with the substrate.

ClpB and the DnaK system cooperate to dissolve protein aggregates in extracts

We tested the capability of cytosolic chaperones to dissolve protein aggregates by using [^{35}S]methionine-labeled soluble extracts of wild-type cells incubated for 15 min at 45°C to induce protein aggregation (see Figure 1). The protein aggregates were isolated by centrifugation, resuspended in buffer with or without combinations of chaperones and ATP, and incubated at permissive temperature (30°C) for 4 h, followed by separation of the soluble and insoluble fractions (Figure 6). No disaggregation occurred in the absence of added chaperones or upon separate addition of the GroEL system, IbpB, HtpG and ClpB. Limited solubilization of only 3–5% of the aggregated protein was observed upon incubation with the DnaK system. In sharp contrast, the simultaneous addition of ClpB and the DnaK system resulted in highly efficient solubilization of ~50% of the aggregated proteins. This disaggregation activity strictly depended on the presence of ATP and was highly efficient since the chaperones were added at substoichiometric amounts (2 μM of ClpB and DnaK monomers) with respect to the proteins (5–10 μM) that aggregated. Furthermore, it exhibited broad substrate specificity as revealed by 2D gel electrophoretic analysis of the soluble and insoluble fractions (Figure 7). Of the aggregated protein species, 20% (50 proteins) were nearly completely disaggregated, 56% (140 proteins) were solubilized by ~50%, and only 24% (60 proteins) were poorly or not disaggregated. We conclude that the DnaK chaperone system cooperates with ClpB to solubilize at least 75% of the thermally aggregated *E.coli* proteins in cell extracts.

Disaggregation of heat-denatured proteins is impaired in dnaK and clpB null mutants

We investigated the potential of ClpB and the DnaK system to reverse protein aggregation *in vivo* by comparing the wild type with isogenic *ΔclpB::kan* and *ΔdnaK52::cat* mutants. Cells of these strains were grown at permissive temperature (30°C) followed by incubation at 45°C for 30 min to induce protein misfolding, and a 3 h recovery period at 30°C to allow protein disaggregation. We noted that the temperature upshift of wild-type cells to 45°C caused a minor and transient accumulation of aggregated proteins, which was not observed after heat shock to 42°C (data not shown). This transient aggregation is consistent with earlier observations (Kucharczyk *et al.*, 1991) and probably results from limitations in chaperone capacity at severe stress conditions (Tomoyasu *et al.*, 1998). Among the transiently aggregated proteins, we identified by mass spectrometry especially large-sized proteins including the β and β′ subunits of RNA polymerase.

In heat-treated *ΔdnaK52::cat* mutant cells, the extensive aggregation of proteins was irreversible (Figure 8), indicating that protein disaggregation depends on DnaK *in vivo*. In heat-treated *ΔclpB::kan* mutant cells, a diversity of proteins aggregated to low amounts that were comparable to those of the transient aggregates observed in wild-type cells (Figure 8). In contrast to wild type, however, these aggregates were not solubilized during the recovery phase. A comparison by 2D gel electrophoresis revealed that the same proteins aggregated transiently in wild-type cells and stably in *ΔclpB::kan*. These proteins form a subset of those proteins that aggregate in *ΔdnaK52::cat* mutant cells (data not shown).

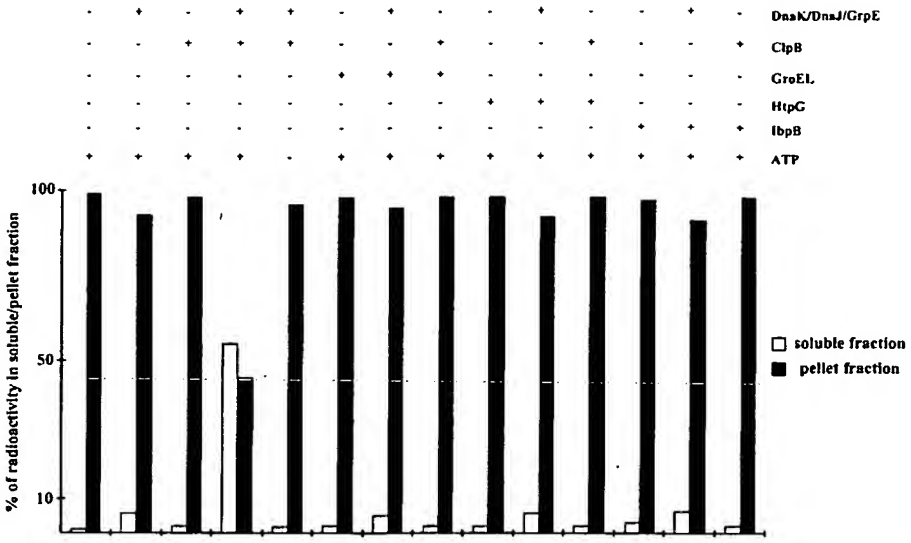


Fig. 6. Disaggregation of pre-existing aggregates of heat-labile *E.coli* proteins in cell extracts. Protein aggregates of [^{35}S]methionine-labeled cell extracts were isolated after heat treatment to 45°C by centrifugation. Pellets were incubated with 2 μM of the indicated chaperone combinations in the presence of 5 mM ATP and an ATP-regenerating system for 4 h at 30°C. Resolubilized protein aggregates and still insoluble protein material were separated by centrifugation. The amount of radioactivity of both fractions (soluble fraction, white bars; pellet fraction, black bars) was determined by scintillation counting. The amount of radioactivity in the control reaction (no chaperone added) before the disaggregation reaction was set at 100%. All reactions were carried out three times with <5% deviations.

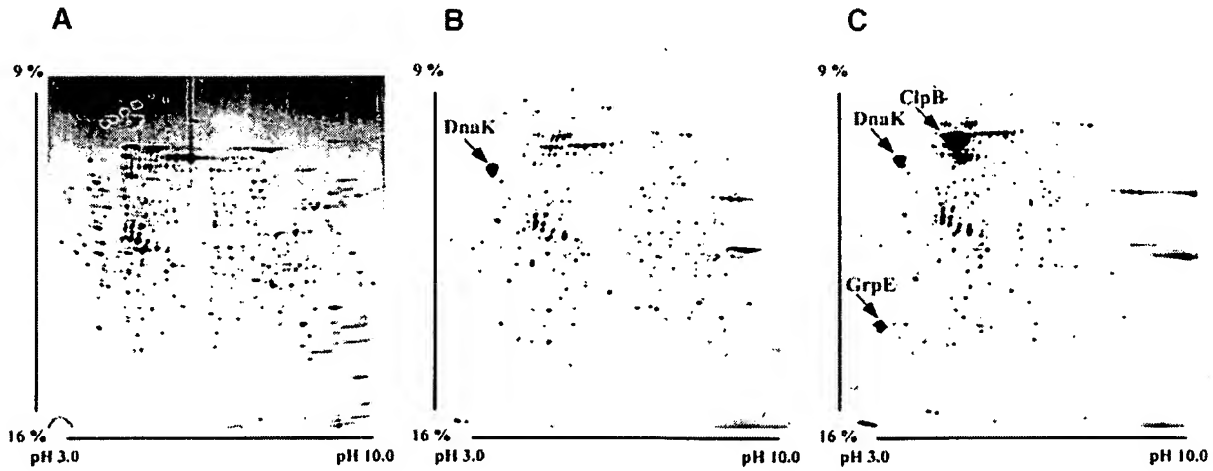


Fig. 7. ClpB and the DnaK system resolubilize pre-existing aggregates of heat-denatured proteins with a broad specificity. Cell extracts were heated to 45°C. Heat-aggregated proteins were isolated by centrifugation and incubated for 4 h at 30°C in the presence of ClpB, the DnaK system and an ATP-regenerating system. Resolubilized and insoluble protein fractions were separated by centrifugation and analyzed by 2D gel electrophoresis. (**A**) Heat-aggregated proteins of *E.coli* cell extract. (**B**) Insoluble and (**C**) soluble (representing resolubilized protein aggregates) protein fractions after the incubation of heat-aggregated proteins with ClpB and the DnaK system.

Reversal of protein aggregation in ΔdnaK52::cat mutants by induced synthesis of ClpB and the DnaK system

To test rigorously whether the extensive aggregation of proteins in heat-treated Δ*dnaK52::cat* mutant cells can be reversed by ClpB and the DnaK system, we performed an *in vivo* order-of-addition experiment. Expression plasmids for *clpB*, *dnaK* and *dnaJ* in various combinations were constructed to allow tight repression of these genes in the absence of inducer and overexpression in the presence of IPTG. Overexpression of *grpE* was not required for these experiments, since Δ*dnaK52::cat* mutant cells have high endogenous GrpE levels due to defects in negative autoregulation of heat-shock genes (Bukau and Walker, 1990). With the continuous presence of IPTG in the growth medium, cells carrying appropriate expression plasmids had levels of ClpB, DnaK and DnaJ that were 5- to 10-fold higher than in wild type. The induced expression of *dnaK* and *dnaJ* without *clpB* (pBB525), or *dnaK* and *dnaJ* with *clpB* (pBB526), rescued the temperature-sensitive growth phenotype of Δ*dnaK52::cat* mutants at 42°C (Table II) and prevented protein aggregation in Δ*dnaK52::cat* mutant cells after the shift from 30 to 42°C (data not shown).

To test for protein disaggregation, Δ*dnaK52::cat* mutant cells carrying appropriate expression plasmids were grown at 30°C in the absence of IPTG, followed by incubation for 30 min at 45°C to allow extensive formation of protein

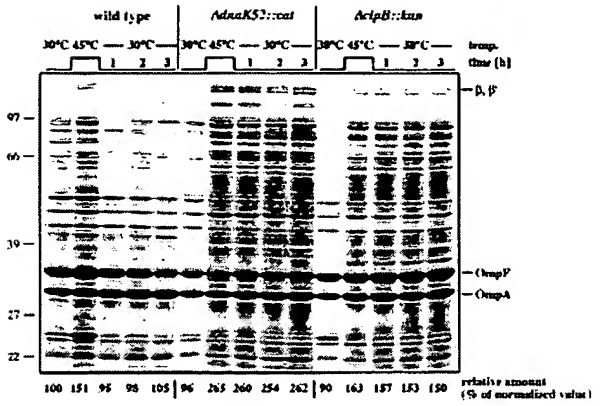


Fig. 8. Disaggregation of heat-induced protein aggregates is impaired in *ΔdnaK52::cat* and *ΔclpB::kan* mutants. Cultures of wild-type and mutant strains were grown in LB medium at 30°C to logarithmic phase, then shifted to 45°C for 30 min, followed by incubation at 30°C for 3 h. Insoluble fractions, containing membrane proteins and aggregated proteins, were analyzed at the indicated time points by SDS–PAGE followed by staining with Coomassie Brilliant Blue. Selected proteins were identified by immunoblot analysis using specific antisera. The insoluble cell fractions were quantified by MacBAS software (Fuji film) and normalized to the band intensities of OmpA and OmpF, which serve as internal loading controls. The amount of insoluble cell fraction of wild-type cells at 30°C before heat shock was set at 100%.

aggregates (Figure 9). After temperature downshift to 30°C, the cells were incubated with or without IPTG and assayed for protein disaggregation. Cells carrying either the vector control (pBB527) or pBB521 (*clpB*) showed no alteration of the high aggregation level in the presence or absence of IPTG (data not shown). Furthermore, even massive IPTG-induced overproduction of other chaperones (GroEL/GroES, HtpG, IbpA/B) did not disaggregate heat-aggregated proteins in *ΔdnaK52::cat* cells (data not shown). In contrast, cells carrying pBB525 (*dnaK dnaJ*) could solubilize ~50% of the pre-existing aggregates when IPTG was present (data not shown). Thus, some solubilization occurs by the DnaK system, probably together with ClpB produced from the chromosomal *clpB* gene. It should be noted that in the *ΔdnaK52::cat* cells, ClpB levels were increased compared with the wild type due to the defects in regulation of heat-shock genes including *clpB* (Bukau and Walker, 1990). The strongest effect was observed for *ΔdnaK52::cat* cells overexpressing *clpB*, *dnaK*, *dnaJ* from pBB526. At least 95% of the heat-aggregated protein species were disaggregated within 2 h at 30°C, provided that IPTG was present after temperature downshift (Figure 9). Unpublished experiments showed that overproduction of the DnaK system in *ΔrpoH* cells, which lack σ^{32} and therefore have strongly reduced levels of all major cytosolic chaperones including ClpB, did not allow solubilization of protein aggregates formed at 42°C. In these cells, co-overproduction of ClpB and the DnaK system was necessary and sufficient to allow complete disaggregation of heat-aggregated proteins (T.Tomoyasu, A.Mogk and B.Bukau, unpublished results). Together, these results demonstrate the generality and high efficiency of the protein disaggregating activity of ClpB and the DnaK system in *E.coli* cells.

MetE: destiny of a thermolabile model protein of E.coli

To strengthen the results obtained in this study and to dissect further the fate of thermolabile proteins, we investigated the most prominent heat-aggregated protein of *ΔdnaK52::cat* cells grown in M9 minimal medium, MetE. Immunoblot analysis revealed that MetE transiently aggregated in wild-type cells subjected to a heat shock to 45°C, but stably aggregated in heat-treated *ΔdnaK52::cat* and *ΔclpB::kan* mutant cells (Figure 10A). Temperature-dependent aggregation and protection by the DnaK system were confirmed in cell extracts and in an *in vitro* system with purified MetE (Figure 10B and C), demonstrating the direct association of the DnaK system with MetE. Efficient resolubilization of pre-formed MetE aggregates by ClpB and the DnaK system was reached in cell extracts and *in vitro* with purified components (Figure 10B and C). These findings confirm for a major natural thermolabile substrate identified in this study, that the DnaK system and ClpB are directly involved in aggregation prevention and disaggregation after temperature-induced unfolding.

Interestingly, MetE was degraded in heat-treated wild-type cells if they were kept at 45°C. This degradation was Lon dependent since MetE was largely stable at 45°C in a *lon* null mutant (Figure 10A), but not in *clpA* or *clpB* null mutants (data not shown). This finding indicates that aggregation-prone substrates that are kept or rendered soluble by the DnaK system may become susceptible to proteolysis.

Discussion

This study shows that the DnaK system is the central 'holder' chaperone of the cytosol that prevents aggregation of a wide variety of heat-denatured proteins. This central function of DnaK is also reflected by its substantially higher effective cellular concentration compared with the other major cytosolic chaperones. We identified for the first time thermolabile proteins of *E.coli* that misfold and aggregate under physiological heat stress. Almost all of them become protected from aggregation by the DnaK system, in at least many cases through direct association with DnaK, which defines them as natural DnaK substrates. Finally, a bi-chaperone system is described consisting of the DnaK system and ClpB, which efficiently disaggregates proteins that escaped the holder function of DnaK.

The thermolabile proteins were identified in heat-treated extracts and *in vivo*. Both populations were qualitatively highly similar, thus validating the *in vitro* experiments, but the extent of aggregation differed strongly between extracts and cells. An identical heat treatment resulted in stable aggregation of 10–15% of the proteins in wild-type extracts, but only transient aggregation of ≤3% of the soluble proteins in wild-type cells. This difference is unlikely to result from the buffer conditions or the chaperone composition in the extracts since they were chosen to be near physiological. We considered that the crowded environment of the *E.coli* cytosol (~170 mg/ml *in vivo* versus 4 mg/ml in the extract) protects proteins from aggregation. A commonly used unspecific crowding agent, Ficoll 70, did not provide protection to proteins against aggregation in cell extracts, even when added to a high concentration of 20% (w/v) (Figure 1). We

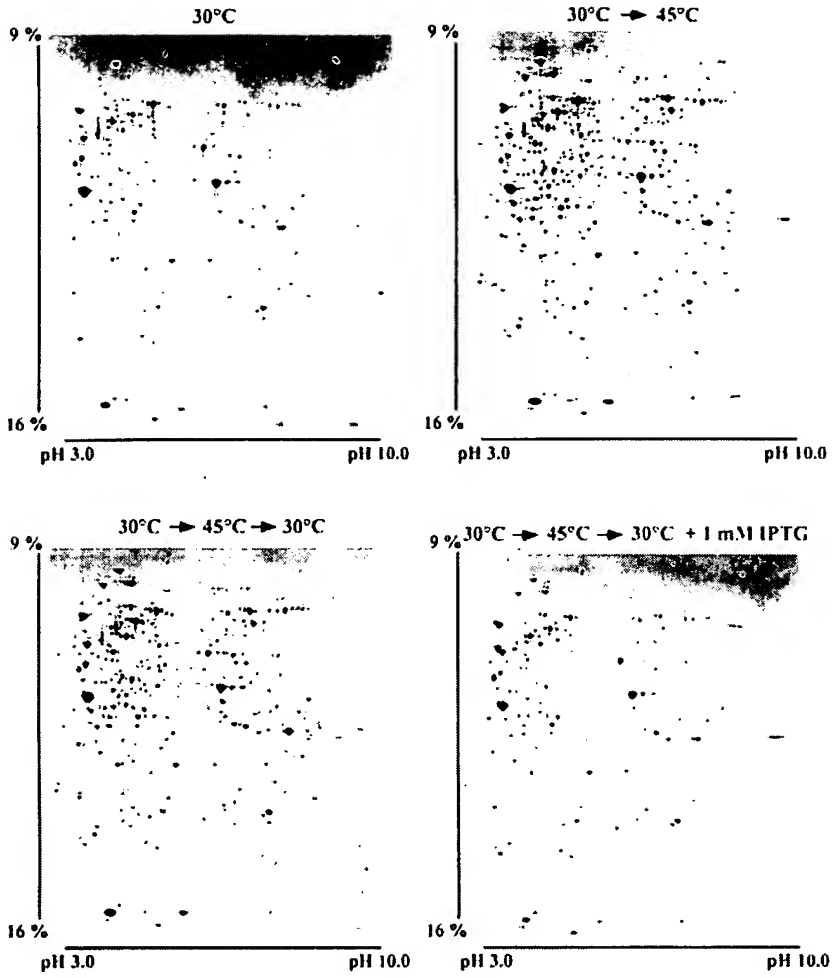


Fig. 9. Disaggregation of heat-induced protein aggregates in *ΔdnaK52::cat* mutant cells by ClpB and the DnaK chaperone machinery. *ΔdnaK52::cat* mutant cells carrying pBB526 (allowing overexpression of *clpB*, *dnaK*, *dnaJ*) were grown in LB medium at 30°C to logarithmic phase and shifted for 30 min to 45°C. Cultures were split after heat treatment and further incubated for 2 h at 30°C with or without IPTG (1 mM). Insoluble cell fractions containing membrane proteins and aggregated proteins were prepared and analyzed by 2D gel electrophoresis.

therefore assume that more specific interactions of the aggregation-prone proteins with other cytosolic chaperones and protectants are responsible for the reduced protein aggregation in wild-type cells compared with cell extracts.

The unique ability of the DnaK system to prevent aggregation of a large variety of thermolabile proteins efficiently was detected *in vivo* and in cell extracts. Incubation of *ΔdnaK52::cat* mutant cells at 42°C caused aggregation of ~10% of the soluble proteins after 60 min. Incubation at 45°C increased the aggregated fraction of each thermolabile protein, but did not change the spectrum of aggregated proteins. The missing function of the DnaK system can thus not be replaced *in vivo* by other chaperones, even though their levels are 2- to 3-fold increased in *ΔdnaK52::cat* mutant cells compared with wild type (Bukau and Walker, 1990). This interpretation is further supported by the findings that IPTG-controlled overproduction of GroEL/GroES, ClpB or IbpA/B did not rescue the temperature-sensitive growth phenotype of *ΔdnaK52::cat* mutant cells at 42°C, and that mutations in the genes encoding these chaperones and HtpG did not cause strong protein aggregation at 42°C. Finally, the experiments using cell extracts showed that only the DnaK system efficiently prevented heat-induced protein aggregation.

This low capacity as holder chaperone was expected for ClpB, since, like its yeast homolog Hsp104, it fails to prevent aggregation of heat-denatured test substrates (Glover and Lindquist, 1998; Goloubinoff *et al.*, 1999). For HtpG our findings agree with earlier reports that HtpG is poorly active in preventing aggregation of unfolded citrate synthase (Jakob *et al.*, 1995). The low capacity of IbpB in cell extracts was unexpected given that stoichiometric amounts prevented aggregation of heat-denatured malate dehydrogenase (Veinger *et al.*, 1998; Shearstone and Baneyx, 1999). However, our findings are consistent with the observed lack of protein aggregation phenotypes in heat-treated *ΔibpAB* null mutants (this study; Thomas and Baneyx, 1998). The low capacity of the GroEL system may be explained by the ability of GroES and ATP to disfavor the GroEL-bound state of thermolabile proteins. Also, the tetradecameric state of the active form of GroEL would demand a 14-fold molar excess of GroEL monomers over the aggregation-prone proteins to reach equimolarity. Such a high GroEL concentration is unphysiological (~3 μM at 30°C), but strong overproduction of GroEL

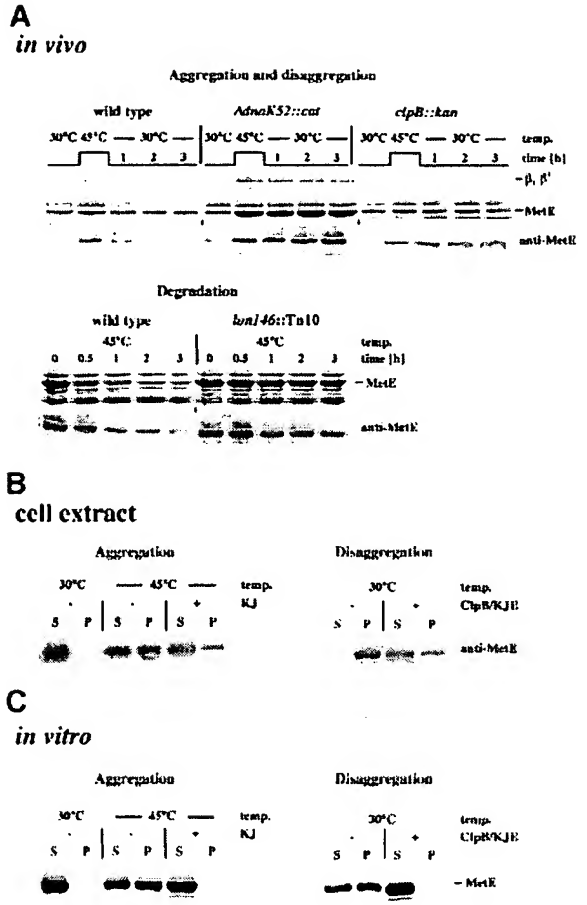


Fig. 10. Aggregation, disaggregation and degradation of MetE. (A) Disaggregation of heat-aggregated MetE is impaired in *ΔdnaK52::cat* and *ΔclpB::kan* mutants. Cultures of wild-type and mutant strains were grown in M9 minimal medium supplemented with 19 L-amino acids except L-methionine at 30°C to logarithmic phase, then shifted to 45°C for 30 min, followed by incubation at 30°C for 3 h. Insoluble fractions, containing membrane proteins and aggregated proteins, were analyzed at the indicated time points by SDS–PAGE followed by staining with Coomassie Brilliant Blue. Transient or stable aggregation of MetE was verified by immunoblot analysis using MetE-specific antisera. MetE is degraded by Lon after heat shock to 45°C in wild-type cells. Cultures of wild-type and *lon146*::Tn*10* mutant strains were grown in supplemented M9 minimal medium at 30°C to logarithmic phase, then shifted to 45°C for 3 h. Soluble fractions were analyzed at the indicated time points by SDS–PAGE followed by staining with Coomassie Brilliant Blue or by immunoblot analysis using MetE-specific antisera. (B) Aggregation and disaggregation of MetE in cell extracts. Cell extracts (4 mg/ml) were heated to 45°C for 15 min in the absence or presence of added 5 μM DnaK and 1 μM DnaJ. Soluble (S) and insoluble (P) fractions were separated by centrifugation and analyzed by immunoblotting using MetE-specific antisera. In addition, heat-aggregated proteins were isolated by centrifugation and incubated with 2 μM ClpB and DnaK, 0.4 μM DnaJ and 0.2 μM GrpE in the presence of 5 mM ATP and an ATP-regenerating system for 4 h at 30°C. Resolubilized and insoluble MetE were separated by centrifugation and analyzed by immunoblotting using MetE-specific antisera. (C) Aggregation and disaggregation of purified MetE *in vitro*. MetE (100 nM) was incubated at 30 or 45°C for 15 min in the absence or presence of DnaK (1 μM) and DnaJ (0.2 μM). Soluble (S) and insoluble (P) fractions were separated by centrifugation and analyzed by silver staining. In addition, ClpB (100 nM) and DnaK/DnaJ/GrpE (100/20/10 nM) were added after heat treatment of MetE and incubated for 4 h at 30°C. Soluble (S) and insoluble (P) fractions were separated by centrifugation and analyzed by silver staining.

and GroES can indeed suppress protein aggregation in chaperone-deficient *ΔrpoH* or *dnaK756* mutant cells (Gragerov *et al.*, 1992; Kedzierska *et al.*, 1999). On the other hand, overproduction of GroEL and GroES could not complement the temperature-sensitive growth phenotype of *ΔdnaK52::cat* mutant cells at 42°C, and only partially suppressed protein aggregation in *ΔdnaK52::cat* at 42°C (this study). These findings confirm the unique role of DnaK as the central protection system of *E.coli* at high temperatures and support results from previous studies (Kusukawa and Yura, 1988). Previously published data (Horwich *et al.*, 1993; Ewalt *et al.*, 1997; Bukau and Horwich, 1998), together with ours, strongly suggest that the primary role of the GroEL system is to promote folding and refolding, rather than holding, of newly synthesized and misfolded proteins.

The number of aggregation-prone *E.coli* proteins is surprisingly high (150–200 *in vivo* and 250 in extracts, out of 800 detectable proteins) given that the chosen heat-shock temperatures are within the growth temperature range. Most aggregation-prone proteins are substrates for DnaK since: (i) addition of DnaK protects their aggregation in cell extracts and *in vivo*; (ii) a highly similar set of proteins aggregates in heat-treated *ΔdnaK52::cat* mutant cells; and (iii) many of these proteins can be co-immunoprecipitated with DnaK-specific antisera in heat-treated cell extracts and *in vivo*. Furthermore, for a major aggregation-prone protein identified *in vivo* and in cell extract, MetE, we showed a direct role for DnaK in aggregation prevention. Purified MetE was found to be thermolabile and prone to aggregation at 45°C *in vitro* and to become protected from aggregation by addition of DnaK and DnaJ. Interestingly, at heat-shock temperatures, MetE aggregated in *ΔdnaK52::cat* mutant cells but was degraded in wild-type cells in a Lon-dependent manner. This example indicates that the fate of thermolabile proteins that were prevented from aggregation by the DnaK system or resolubilized by ClpB and the DnaK system (see below) can also be to be targeted for degradation. Degradation is likely to explain why the patterns of DnaK substrates identified *in vivo* and *in vitro* show some differences.

All identified proteins are cytoplasmic and participate in various cellular processes including metabolism, transcription, translation, cell wall synthesis and cell division, and in particular include key proteins of the transcription and translation apparatus. Many of these proteins have essential functions, and aggregation of any one of them may cause the temperature sensitivity of *ΔdnaK52::cat* mutants. The thermosensitivity of key proteins of transcription (RNA polymerase) and translation (elongation factors) may have protective roles in heat-treated cells, since it may slow down protein synthesis when the stress state limits the availability of chaperones. The insoluble fraction of heat-treated *ΔdnaK52::cat* mutant cells and cell extracts also contained the proteases Lon, ClpX and HslU, as well as IbpB (not in cell extracts). These components of the cellular folding system may not be thermolabile but rather associate with denatured and aggregated protein substrates as already reported for IbpB (Allen *et al.*, 1992; Laskowska *et al.*, 1996b). In isolated cases proteins may become trapped in aggregates by specific association with a thermolabile protein as part of

an oligomeric complex (e.g. Rho). We consider this possibility as an exception, since most aggregation-prone proteins that we identified participate in different cellular processes. Furthermore, we have excluded experimentally the possibility that the aggregation of thermolabile proteins causes unspecific co-precipitation of stable proteins.

Earlier attempts to identify the parameters determining the thermolability of model proteins gave no generally applicable answer (Jaenicke and Bohm, 1998). Our study provides a different access to this problem by identifying a large number of aggregation-prone proteins of *E.coli* and showing that they are substrates for DnaK upon thermal unfolding. The most striking feature of this class is that large-sized proteins are strongly enriched. It is tempting to speculate that the appearance of large, multi-domain proteins in evolution was accompanied by the appearance of a powerful DnaK chaperone system. Three features of large proteins may contribute to their aggregation propensity. First, misfolded conformers of larger proteins may statistically expose more hydrophobic surface patches than smaller proteins. Since such patches are considered to be involved in, or even trigger, intermolecular aggregation (Mitraki and King, 1989; Seckler and Jaenicke, 1992), there is a higher probability that larger proteins will aggregate. Secondly, large proteins are composed of more domains than small proteins. In many proteins the interactions between domains are flexible and subject to regulation. Interdomain surface contacts may be particularly vulnerable to heat and, if they expose hydrophobic surfaces, initiate intermolecular aggregation. Thirdly, the rates of refolding of unfolded conformers of large proteins may be slower than that of small proteins, which consequently favors competing aggregation reactions.

All features involve the surface exposure of hydrophobic patches, which may be the connection between intermolecular aggregation and DnaK association. We previously established that DnaK recognizes hydrophobic patches consisting of 4–5 consecutive residues enriched in large hydrophobic and aromatic residues and flanked by regions enriched in basic residues (Rüdiger *et al.*, 1997a,b). An algorithm predicting DnaK-binding sites in protein sequences was established (Rüdiger *et al.*, 1997b) and used here to compare 32 thermolabile proteins shown to be DnaK substrates by co-immunoprecipitation with DnaK-specific antisera and/or shown to become aggregated or disaggregated in a DnaK-dependent fashion with 27 thermostable control proteins. The majority of the thermolabile and stable proteins have statistically one DnaK-binding site every 26 and 35 residues, respectively. Nineteen percent of the thermolabile subset have a higher frequency (compared with 4% of the thermostable subset), while 41% of the stable subset have a lower frequency (compared with 16% of the thermolabile subset). There is thus some correlation between the frequency of predicted DnaK-binding sites and the propensity to aggregate or associate with DnaK. It is interesting that of all *E.coli* proteins >700 residues investigated, only FumC (1010 residues) was stable in *ΔdnaK52::cat* mutants at 42°C, the protein that has the lowest frequency of DnaK-binding sites of all proteins investigated (one binding site per 92 residues).

One can not predict the particular hydrophobic patch(es) of a thermolabile protein involved in aggregation and DnaK binding, given that aggregation is primarily driven by unfolding kinetics and properties of the unfolded state. We nevertheless investigated whether thermolabile DnaK substrates share features of their native structures, especially regarding the localization of potential DnaK-binding sites. We compared the atomic structures of DnaK substrates (GlnA, GatY, EF-G, AsnA, DeoA, CarA, CarB, SerA and GuaA) with those of stable proteins (DeoD, FabF, PckA, FbaB, EF-Ts, PykF, GlpK and FumC). In eight out of nine thermolabile proteins analyzed (all except DeoA), at least one patch of three hydrophobic side chains, constituting the core of a DnaK-binding site, is surface exposed, and in three of these nine proteins several such patches are exposed [SerA and GuaA (Figure 5B); complex of CarA and CarB (not shown)]. In contrast, only one of the eight stable proteins exposes such a hydrophobic patch (FbaB). It should be emphasized that the mere surface exposure of these sites in the native structures is insufficient for DnaK association. The architecture of the substrate-binding cavity of DnaK indeed requires significant separation of the binding segment from the remainder of the polypeptide chain (Zhu *et al.*, 1996). One possible scenario may therefore be that these surface-exposed sites become even more exposed upon thermal unfolding of the proteins and trigger protein aggregation. Furthermore, we studied the contact sites between protein domains and subunits and found potential DnaK-binding sites on interdomain contact surfaces, particularly of smaller domains (Figure 5C) and some inter-subunit surfaces (Figure 5D). Although these features are not unique for the proteins of the subset of thermolabile proteins, the accessibility of such contact sites may be increased during thermal unfolding in the case of thermolabile proteins.

Our study provides evidence for a highly efficient bi-chaperone system, composed of the DnaK system and ClpB, for dissolving a wide spectrum of aggregated proteins in extracts and *in vivo*. These results extend the pioneering findings of Lindquist and coworkers (Sanchez *et al.*, 1992; Parsell *et al.*, 1994; Glover and Lindquist, 1998), which for the first time established a protein disaggregating activity for the yeast homologs Hsp104 and Ssa1/Ydj1. ClpB and the DnaK system are both required for this activity and can not be replaced by any of the other chaperones tested. Substoichiometric amounts of both chaperone systems (2 μM for ClpB and DnaK) were sufficient to solubilize ~50% of the amount of aggregated proteins (5–10 μM) present in heat-treated cell extracts, with at least partial solubilization of 75% of the aggregated protein species. These findings show the high efficiency and broad specificity of this bi-chaperone system. The efficiency was even higher *in vivo* than *in vitro*, as it allowed complete solubilization of pre-formed, large-sized aggregates. A central role of this bi-chaperone system in stress protection of *E.coli* is indicated by the failure of *ΔdnaK52::cat* and *ΔclpB::kan* mutant cells to disaggregate protein aggregates formed during severe heat treatment. In agreement with our data, *clpB* null mutants and *dnaK* point mutants are retarded in removal of a cell fraction assumed to be enriched in protein aggregates (Laskowska *et al.*, 1996a; Kedzierska *et al.*, 1999).

Our findings functionally explain genetic evidence that suggests a stress-related role of the cooperative activity

of the DnaK system and ClpB. First, in several bacteria the heat-shock-inducible *clpB*, *dnaK* and *dnaJ* genes are organized in an operon (Falah and Gupta, 1997; Osipiuk and Joachimiak, 1997), or are at least regulated by a common, specific mechanism (Grandvalet *et al.*, 1999). Secondly, knockout mutations in *clpB* of *E.coli* and *Helicobacter pylori*, and in *hsp104* of *Saccharomyces cerevisiae* prevent the cells from acquiring thermotolerance and surviving severe stress conditions (Lindquist, 1990; Squires *et al.*, 1991; Sanchez *et al.*, 1992; Lindquist and Kim, 1996; Allan *et al.*, 1998; Thomas and Baneyx, 1998). Thermotolerance and survival of bacteria and yeast under heat stress are thus linked to the ability of these cells to reverse protein aggregation, an activity that is performed by the bi-chaperone system.

Materials and methods

Strains and culture conditions

Escherichia coli strains used were derivatives of MC4100 [araD139 *Δ(argF-lac)U169 rpsL150 relA1 flbB5301 deoC1 ptsF25 rbsR*]. Mutant alleles of chaperone genes were introduced by P1 transduction into MC4100 background to generate strains BB1553 (*ΔdnaK52::cat sidB1*) (Bukau and Walker, 1990), BB2395 (*lon*146::Tn*10*), BB2396 (*clpA*82::Tn*10*) (Gottesman and Maurizi, 1992), BB4561 (*ΔclpB::kan*) (Squires *et al.*, 1991), BB4562 (*ΔhtpG1::lacZ zba315::kan*) (Bardwell and Craig, 1988), BB4563 (*ΔibpAB::kan*) (S.Vorderwülbecke, E.Deuerling and B.Bukau, unpublished), BB4564 (*groEL140 zjd::*Tn*10 zje::*ΩSpc^r/Str^r) and BB4565 (*groEL44 zjd::*Tn*10 zje::kan*) (Zeilstra-Ryalls *et al.*, 1993). Bacterial strains were cultured at 30°C in Luria broth (LB) or M9 minimal medium supplemented with 0.2% (w/v) glucose and 19 L-amino acids (Sigma LAA-21; 50 μg/ml/amino acid). Chloramphenicol, kanamycin, spectinomycin, tetracycline and ampicillin were used at final concentrations of 20, 20, 50, 10 and 100 μg/ml, respectively. Temperature-shift experiments were performed in orbital shaking water baths. For pulse-labeling with [^{35}S]methionine, cells were grown in M9/glucose with all L-amino acids except L-methionine. Labeling was done by adding [^{35}S]methionine (Amersham SJ1515; 15 mCi/ml, 1000 Ci/mmol) to 30 μCi/ml cell culture for 2 min, followed by addition of unlabeled L-methionine to 200 mg/ml. In co-immunoprecipitation experiments, cells were pulse-labeled three times to increase the labeling efficiency of the proteins.

Protein purifications

Chaperones were purified according to standard protocols (Jakob *et al.*, 1995; Buchberger *et al.*, 1996; Kim *et al.*, 1998; Veinger *et al.*, 1998). The DnaK and GroEL systems, IbpB and HtpG, were active in preventing aggregation and assisting refolding of denatured model substrates (malate dehydrogenase, firefly luciferase, mitochondrial citrate synthase) similar to published efficiencies (Jakob *et al.*, 1995; Buchberger *et al.*, 1996; Veinger *et al.*, 1998). MetE was purified following published protocols (Gonzalez *et al.*, 1992).

Preparation of soluble cell extracts and spheroplasts

Cell cultures were grown to mid-exponential phase and cooled rapidly in an ice-water bath to 0°C. Cells were harvested by centrifugation for 10 min at 5000 *g* and 4°C. Cell pellets were washed twice and resuspended in ice-cold breakage buffer (50 mM HEPES pH 7.6, 150 mM KCl, 20 mM $MgCl_2$, 10 mM DTT) and lysed in a precooled French pressure cell at 20 000 lb/in^2. Total cell protein was quantified by Bradford assay using BSA as standard. Spheroplasts were prepared following published protocols (Ewalt *et al.*, 1997). For co-immunoprecipitations spheroplasts were resuspended in M9/glucose medium containing all L-amino acids except L-methionine and 15% (w/v) sucrose.

Aggregation and disaggregation of proteins in cell extracts

Soluble cell extracts (4 mg/ml) were pre-incubated for 5 min at 30°C in the absence or presence of chaperones. ATP (10 mM) was added 2 min prior to heat shock. Samples were shifted to 45°C for 15 min and aggregated proteins were pelleted by centrifugation for 15 min at 15 000 *g* and 4°C. Pellets were washed twice with ice-cold breakage buffer and analyzed by electrophoresis. For disaggregation of aggregated proteins the washed protein pellet was resuspended in breakage buffer by pipetting. Chaperones (2 μM DnaK, ClpB; 0.4 μM DnaJ; 0.2 μM GrpE), ATP (5 mM) and an ATP-regenerating system (3 mM phospho-enolpyruvate, 20 ng/ml pyruvate kinase) were added and incubated for 4 h at 30°C. Resolubilized and still aggregated proteins were separated by centrifugation (15 min, 15 000 *g* at 4°C). Pellets were washed twice with ice-cold breakage buffer and the soluble and insoluble fractions were further analyzed by gel electrophoresis.

Preparation of insoluble cell fractions

Aliquots (10–40 ml) of bacterial cultures were cooled rapidly to 0°C in an ice-water bath and within a few minutes centrifuged for 10 min at 5000 *g* and 4°C to harvest the cells. Pellets were resuspended in 10× lysis buffer [100 mM Tris pH 7.5, 100 mM KCl, 2 mM EDTA, 15% (w/v) sucrose, 1 mg/ml lysozyme] according to their optical density (50 ml of lysis buffer for 10 ml of culture of OD_{600} = 1) and frozen at –80°C. Samples were then thawed slowly at 0°C followed by addition of a 10-fold volume of ice-cold water and mixing. The viscous, turbid solution was sonicated with a Branson Cell Disruptor B15 (microtip, level 5, 50% duty, four cycles) while cooling. Intact cells were removed by centrifugation at 2000 *g* for 15 min at 4°C. The insoluble cell fraction was isolated by subsequent centrifugation at 15 000 *g* for 15 min at 4°C. Pellets were washed twice with ice-cold water and analyzed by gel electrophoresis.

Co-immunoprecipitation

[^{35}S]methionine-labeled soluble cell extracts (4 mg/ml) were incubated at 30 or 45°C in the presence of added DnaK, DnaJ and GrpE (7.5, 1.5 and 0.75 μM, respectively). After incubation for 15 min at the temperature indicated, ice-cold EDTA (100 mM) was added and the samples were kept on ice to disfavor dissociation of DnaK from the substrates. [^{35}S]methionine-labeled spheroplasts of MC4100 cells were incubated at 30 or 45°C for 5 min, lysed by dilution into 10 vols of ice-cold EDTA (100 mM), followed by sonication. Intact cells and membrane fractions were removed by subsequent centrifugation at 15 000 *g* for 15 min at 4°C. Aliquots of 80 μl were incubated with 50 μl of protein A–Sepharose beads in phosphate-buffered saline (PBS), 4 μl of polyclonal antisera raised against the ATPase domain of DnaK, and 1 ml of co-immunoprecipitation (CoIP) buffer [50 mM Tris pH 7.5, 10 mM EDTA, 150 mM KCl, 0.5% (v/v) NP-40] for 1 h at 4°C. Beads were washed twice with 1 ml of CoIP buffer and once with 1 ml of PBS containing 5 mM EDTA. Proteins bound to protein A–Sepharose were analyzed by gel electrophoresis. Dried gels were scanned using a phosphoimager (FLA-2000) and quantified by MacBAS software (Fuji film).

Two-dimensional gel electrophoresis

Two-dimensional gel electrophoresis was performed essentially as described previously (Bjellqvist *et al.*, 1993) with minor modifications. Nonlinear immobilized pH 3–10 gradient strips (18 cm; Pharmacia Biotech) were rehydrated in 7 M urea, 2 M thiourea, 2% (w/v) CHAPS, 0.4% (w/v) DTT, 2% (v/v) resolyte pH 3–10 (Pharmacia Biotech). Total soluble cell extracts were dialyzed twice against 2 l of 5 mM HEPES pH 8.0, containing 10% (v/v) glycerol and 1 mM EDTA. After dialysis, urea (8 M), CHAPS [2% (w/v)], DTT [0.4% (w/v)] and 0.5% (v/v) resolyte pH 3–10 were added and ~1 mg of total soluble protein was applied to both the cathodic and anodic ends of the rehydrated strips. In the case of insoluble cell fractions the pellets were dissolved in 7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 0.4% (w/v) DTT and 0.5% (v/v) resolyte pH 3–10. The proteins were focused first for 2 h at 150 V, then for 5 h at 300 V. The voltage was finally increased to 1500 V for 10 h and to 3500 V for 48–60 h. The proteins were separated on 9–16% linear gradient polyacrylamide gels at 5 mA/gel for 2 h, followed by 40 mA/gel for 7 h. The gels were stained with colloidal Coomassie Blue (Novex) and destained with water for protein detection.

Identification of proteins by mass spectrometry

For protein identification the Coomassie-stained spots were cut out of the gel and destained by a solution of 100 mM ammonium bicarbonate in 50% acetonitrile/water. The gel pieces were then prepared and digested with endoproteinase Lys-C essentially as described previously (Langen *et al.*, 1997). Elution of the resulting peptides and determination of their masses followed published protocols (Fountoulakis and Langen, 1997). Proteins were also identified by protein spot matching with reference gels using the ImageMaster software (Pharmacia).

SDS–PAGE, immunoblotting and quantifications

Gel electrophoresis was carried out according to published protocols (Laemmli, 1970) using 12% SDS–PAGE and stained with Coomassie

Brilliant Blue or silver nitrate. Immunoblotting was performed according to standard procedures, using rabbit antisera specific for the relevant protein as primary antibody, and developed with a Vistra ECF fluorescence immunoblotting kit (Amersham) using alkaline phosphatase/conjugated anti-rabbit IgG as secondary antibodies (Vector Laboratories). Developed immunoblots were scanned using a fluoroimager (FLA-2000) and quantified using MacBAS software (Fuji Film). To determine the cellular levels of relevant proteins, different amounts of total and soluble cell extracts were subjected to SDS–PAGE. For immunoblot quantification and determination of the linear range of detection, serial dilutions of purified proteins (ClpB, HtpG, DnaK, GroEL, IbpB) served as an internal standard. Protein molecules/cell were calculated on the basis of the known weight of proteins per cell (0.154 pg) (Neidhardt and VanBogelen, 1987).

Construction of plasmids

Control plasmid pBB527 (pZE4 *lacI*) was derived from pZE GrpE (laboratory collection, derivative of pUHS4 IntI) (Lutz and Bujard, 1997). pZE GrpE and pBB509 (Tomoyasu *et al.*, 1998) were digested by *Avr*II or *Eco*RI and the sticky ends created were filled in with T4 DNA polymerase (Boehringer Mannheim). Both plasmids were then digested with *Aat*II and the isolated *lacI* fragment was subsequently cloned into engineered pZE pzl GrpE. pBB525 (pZE4 *lacI* pA1/lacO1 *dnaK*, *dnaJ*) was derived from pBB518 (pZE1 *lacI* pA1/lacO1 *dnaK*, *dnaJ*; laboratory collection). pBB518 and pZE GrpE were digested with *Aat*II and *Avr*II. The isolated *lacI*, *dnaK*, *dnaJ* containing fragment was ligated into engineered pZE GrpE. The *clpB* gene from *E.coli* was amplified by PCR using native Pfu DNA polymerase (Stratagene), pClpB (Squires *et al.*, 1991) as template, forward primer 5′-AGGAAGATCTATGGGAGGAGTTATGCGTCT-3′ and reverse primer 5′-GGCCTCTAGATTACTGGACGGCGACAATCC-3′. The engineered *Bgl*II and *Xba*I sites were used for subsequent cloning of the *clpB* gene fragment into the single *Bam*HI and *Xba*I restriction sites of pUHE21-2fdΔ12 (Buchberger *et al.*, 1994) followed by sequencing of the entire gene. The resulting plasmid pBB519 (pUHE21-2fdΔ12 *clpB*) was digested with *Xho*I and *Xba*I, and ligated into *Xho*I–*Avr*II-digested pZE GrpE. The isolated *clpB* fragment was subsequently cloned into engineered pZE GrpE, yielding pBB521 (pZE4 pA1/lacO1 *clpB*). pBB526 (pZE4 lacI pA1/lacO1 *dnaK*, *dnaJ*, *clpB*) was constructed by digesting pBB525 with *Avr*II and pBB521 with *Xho*I, and the sticky ends created were filled in with T4 DNA polymerase. Both resulting plasmids were cut with *Aat*II afterwards and the *lacI*, *dnaK*, *dnaJ* containing fragment was cloned into engineered pBB521. The *ibpAB* operon from *E.coli* was amplified by PCR using native Pfu DNA polymerase, chromosomal DNA from *E.coli* C600 as template, forward primer 5′-AGGAGGATCCATGCGTAACTTTGATTTATCC-3′ and reverse primer 5′-AGGAAAGCTTTTAGCTATTTAACGCGGGAC-3′. The engineered *Bam*HI and *Hin*dIII sites were used for subsequent cloning of the *ibpAB* operon into single *Bam*HI and *Hin*dIII restriction sites of pUHE21-2fdΔ12 followed by sequencing of the entire operon. The *groESL* operon from *E.coli* was amplified by PCR using native Pfu DNA polymerase, chromosomal DNA from *E.coli* C600 as template, forward primer 5′-AGGAGTCGACATGAATATTCGTCCATTGCAT-3′ and reverse primer 5′-AGGAGTCGATTTACATCATGCCGCCCATGC-3′. The engineered *Sal*I sites were used for subsequent cloning of the *groESL* operon into the single *Sal*I restriction site of pDS56 followed by sequencing of the entire operon. The *htpG* gene from *E.coli* was amplified by PCR using native Pfu DNA polymerase, chromosomal DNA from *E.coli* C600 as template, forward primer 5′-GGCCATGGATCCATGAAAGGACAAGAAACTCGTGGT-3′ and reverse primer 5′-GGCCATAAGCTTTCAGGAAACCAGCAGCTGGTTCAT-3′. The engineered *Bam*HI and *Hin*dIII sites were used for subsequent cloning of the *htpG* gene into single *Bam*HI and *Hin*dIII restriction sites of pUHE21-2fdΔ12 followed by sequencing of the entire gene.

Acknowledgements

We thank A.Jacobi and M.Mayer for purified chaperones; S.Vorderwülbecke and E.Deuerling for construction of the *ΔibpAB::kan* strain and plasmid pUHE21-2fdΔ12-*ibpAB*; C.Georgopoulos, J.Bardwell, S.Gottesman and C.Squires for strains; E.Deuerling for discussions and comments on the manuscript; and T.Rauch for data evaluation by ImageMaster software (Pharmacia). This work was supported by grants from the Alexander von Humboldt Stiftung to P.G., the Deutsche Forschungsgemeinschaft (SFB388) and the Fonds der Chemie to B.B., and in part by a grant of the Bundesministerium für Bifolung und Forschung to B.B.

References

Allan,E., Mullany,P. and Tabaqchali,S. (1998) Construction and characterization of a *Helicobacter pylori clpB* mutant and role of the gene in the stress response. *J. Bacteriol.*, **180**, 426–429.

Allen,S.P., Polazzi,J.O., Gierse,J.K. and Easton,A.M. (1992) Two novel heat shock genes encoding proteins produced in response to heterologous protein expression in *Escherichia coli*. *J. Bacteriol.*, **174**, 6938–6947.

Bardwell,J.C.A. and Craig,E.A. (1988) Ancient heat shock gene is dispensable [published erratum appears in *J. Bacteriol.*, **170**, 4999]. *J. Bacteriol.*, **170**, 2977–2983.

Bjellqvist,B., Sanchez,J.C., Pasquali,C., Ravier,F., Paquet,N., Frutiger,S., Hughes,G.J. and Hochstrasser,D. (1993) Micropreparative two-dimensional electrophoresis allowing the separation of samples containing milligram amounts of proteins. *Electrophoresis*, **14**, 1375–1378.

Buchberger,A., Valencia,A., McMacken,R., Sander,C. and Bukau,B. (1994) The chaperone function of DnaK requires the coupling of ATPase activity with substrate binding through residue E171. *EMBO J.*, **13**, 1687–1695.

Buchberger,A., Schröder,H., Hesterkamp,T., Schönfeld,H.-J. and Bukau,B. (1996) Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding. *J. Mol. Biol.*, **261**, 328–333.

Bukau,B. (1999) *Molecular Chaperones and Folding Catalysis—Regulation, Cellular Function and Mechanisms*. Harwood Academic Publishers, Amsterdam, The Netherlands.

Bukau,B. and Horwich,A.L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell*, **92**, 351–366.

Bukau,B. and Walker,G. (1990) Mutations altering heat shock specific subunit of RNA polymerase suppress major cellular defects of *E.coli* mutants lacking the DnaK chaperone. *EMBO J.*, **9**, 4027–4036.

Craig,E.A. and Gross,C.A. (1991) Is hsp70 the cellular thermometer? *Trends Biochem. Sci.*, **16**, 135–140.

Deuerling,E., Schulze-Specking,A., Tomoyasu,T., Mogk,A. and Bukau,B. (1999) Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature*, **400**, 693–696.

Ehrnsperger,M., Gräber,S., Gaestel,M. and Buchner,J. (1997) Binding of non-native protein to Hsp25 during heat shock creates a reservoir of folding intermediates for reactivation. *EMBO J.*, **16**, 221–229.

Ewalt,K.L., Hendrick,J.P., Houry,W.A. and Hartl,F.U. (1997) *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell*, **90**, 491–500.

Falah,M. and Gupta,R.S. (1997) Phylogenetic analysis of mycoplasmas based on Hsp70 sequences: cloning of the *dnaK* (*hsp70*) gene region of *Mycoplasma capricolum*. *Int. J. Syst. Bacteriol.*, **47**, 38–45.

Fountoulakis,M. and Langen,H. (1997) Identification of proteins by matrix-assisted laser desorption ionization-mass spectrometry following in-gel digestion in low-salt, nonvolatile buffer and simplified peptide recovery. *Anal. Biochem.*, **250**, 153–156.

Freeman,B.C. and Morimoto,R.I. (1996) The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. *EMBO J.*, **15**, 2969–2979.

Glover,J.R. and Lindquist,S. (1998) Hsp104, Hsp70 and Hsp40: a novel chaperone system that rescues previously aggregated proteins. *Cell*, **94**, 73–82.

Goloubinoff,P., Mogk,A., Peres Ben Zvi,A., Tomoyasu,T. and Bukau,B. (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. *Proc. Natl Acad. Sci. USA*, **96**, 13732–13737.

Gonzalez,J.C., Banerjee,R.V., Huang,S., Sumner,J.S. and Matthews,R.G. (1992) Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: two solutions to the same chemical problem. *Biochemistry*, **31**, 6045–6056.

Gottesman,S. and Maurizi,M.R. (1992) Regulation by proteolysis: energy-dependent proteases and their targets. *Microbiol. Rev.*, **56**, 592–621.

Gragerov,A., Nudler,E., Komissarova,N., Gaitanaris,G., Gottesman,M. and Nikiforov,V. (1992) Cooperation of GroEL/GroES and DnaK/DnaJ heat shock proteins in preventing protein misfolding in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **89**, 10341–10344.

Grandvalet,C., de Crecy-Lagard,V. and Mazodier,P. (1999) The ClpB ATPase of *Streptomyces albus* G belongs to the HspR heat shock regulon. *Mol. Microbiol.*, **31**, 521–532.

Hartl,F.U. (1996) Molecular chaperones in cellular protein folding. *Nature*, **381**, 571–580.

Hesterkamp,T. and Bukau,B. (1998) Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in *E.coli*. *EMBO J.*, **17**, 4818–4828.

Horwich,A.L. and Weissman,J.S. (1997) Deadly conformations—protein misfolding in prion disease. *Cell*, **89**, 495–510.

Horwich,A.L., Low,K.B., Fenton,W.A., Hirshfield,I.N. and Furtak,K. (1993) Folding *in vivo* of bacterial cytoplasmic proteins: role of GroEL. *Cell*, **74**, 909–917.

Hwang,D.S., Crooke,E. and Kornberg,A. (1990) Aggregated dnaA protein is dissociated and activated for DNA replication by phospholipase or DnaK protein. *J. Biol. Chem.*, **265**, 19244–19248.

Jaenicke,R. and Bohm,G. (1998) The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.*, **8**, 738–748.

Jakob,U., Lilie,H., Meyer,I. and Buchner,J. (1995) Transient interaction of Hsp90 with early unfolding intermediates of citrate synthase. Implications for heat shock *in vivo*. *J. Biol. Chem.*, **270**, 7288–7294.

Johnson,J.L. and Craig,E.A. (1997) Protein folding *in vivo*: unraveling complex pathways. *Cell*, **90**, 201–204.

Kedzierska,S., Staniszewska,M., Wegrzyn,A. and Taylor,A. (1999) The role of DnaK/DnaJ and GroEL/GroES systems in the removal of endogenous proteins aggregated by heat-shock from *Escherichia coli* cells. *FEBS Lett.*, **446**, 331–337.

Kim,K.I., Woo,K.M., Seong,I.S., Lee,Z.W., Baek,S.H. and Chung,C.H. (1998) Mutational analysis of the two ATP-binding sites in ClpB, a heat shock protein with protein-activated ATPase activity in *Escherichia coli*. *Biochem. J.*, **333**, 671–676.

Kucharczyk,K., Laskowska,E. and Taylor,A. (1991) Response of *Escherichia coli* cell membranes to induction of λ cI857 prophage by heat shock. *Mol. Microbiol.*, **5**, 2935–2945.

Kusukawa,N. and Yura,T. (1988) Heat shock protein GroE of *Escherichia coli*: key protective roles against thermal stress. *Genes Dev.*, **2**, 874–882.

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

Langen,H., Gray,C., Röder,D., Juranville,J.-F., Takacs,B. and Fountoulakis,M. (1997) From genome to proteome: protein map of *Haemophilus influenzae*. *Electrophoresis*, **18**, 1184–1192.

Langer,T., Pfeifer,G., Martin,J., Baumeister,W. and Hartl,F.-U. (1992) Chaperonin-mediated protein folding: GroES binds to one end of the GroEL cylinder, which accommodates the protein substrate within its central cavity. *EMBO J.*, **11**, 4757–4765.

Laskowska,E., Kuczynska-Wisnik,D., Skórko-Glonek,J. and Taylor,A. (1996a) Degradation by proteases Lon, Clp and HtrA, of *Escherichia coli* proteins aggregated *in vivo* by heat shock; HtrA protease action *in vivo* and *in vitro*. *Mol. Microbiol.*, **22**, 555–571.

Laskowska,E., Wawrzynow,A. and Taylor,A. (1996b) IbpA and IbpB, the new heat-shock proteins, bind to endogenous *Escherichia coli* proteins aggregated intracellularly by heat shock. *Biochimie*, **78**, 117–122.

Lindquist,S. and Kim,G. (1996) Heat-shock protein 104 expression is sufficient for thermotolerance in yeast. *Proc. Natl Acad. Sci. USA*, **93**, 5301–5306.

Lindquist,S. and Schirmer,E.C. (1999) The role of Hsp104 in stress tolerance and prion maintenance. In Bukau,B. (ed.), *Molecular Chaperones and Folding Catalysts—Regulation, Cellular Function and Mechanism*. Harwood Academic Publishers, Amsterdam, The Netherlands, pp. 347–380.

Lutz,R. and Bujard,H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I_1-I_2 regulatory elements. *Nucleic Acids Res.*, **25**, 1203–1210.

Martin,J. and Hartl,F.-U. (1997) The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proc. Natl Acad. Sci. USA*, **94**, 1107–1112.

Mitraki,A. and King,J. (1989) Protein folding intermediates and inclusion body formation. *Biotechnology*, **7**, 690–697.

Morimoto,R., Tissieres,A. and Georgopoulos,C. (1994) *The Biology of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Neidhardt,F.C. and VanBogelen,R.A. (1987) Heat shock response. In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. Vol. 2. American Society for Microbiology, Washington, DC, pp. 1334–1345.

Netzer,W.J. and Hartl,F.U. (1998) Protein folding in the cytosol: chaperonin-dependent and -independent mechanisms. *Trends Biochem. Sci.*, **23**, 68–73.

Osipiuk,J. and Joachimiak,A. (1997) Cloning, sequencing and expression of dnaK-operon proteins from the thermophilic bacterium *Thermus thermophilus*. *Biochim. Biophys. Acta*, **1353**, 253–265.

Parsell,D.A., Kowal,A.S., Singer,M.A. and Lindquist,S. (1994) Protein disaggregation mediated by heat-shock protein Hsp104. *Nature*, **372**, 475–478.

Rüdiger,S., Buchberger,A. and Bukau,B. (1997a) Interaction of Hsp70 chaperones with substrates. *Nature Struct. Biol.*, **4**, 342–349.

Rüdiger,S., Germeroth,L., Schneider-Mergener,J. and Bukau,B. (1997b) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.*, **16**, 1501–1507.

Sanchez,Y. and Lindquist,S.L. (1990) HSP104 required for induced thermotolerance. *Science*, **248**, 1112–1115.

Sanchez,Y., Taulin,J., Borkovich,K.A. and Lindquist,S. (1992) Hsp104 is required for tolerance to many forms of stress. *EMBO J.*, **11**, 2357–2364.

Schröder,H., Langer,T., Hartl,F.-U. and Bukau,B. (1993) DnaK, DnaJ, GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J.*, **12**, 4137–4144.

Seckler,R. and Jaenicke,R. (1992) Protein folding and protein refolding. *FASEB J.*, **6**, 2545–2552.

Shearstone,J.R. and Baneyx,F. (1999) Biochemical characterization of the small heat shock protein IbpB from *Escherichia coli*. *J. Biol. Chem.*, **274**, 9937–9945.

Skowyra,D., Georgopoulos,C. and Zylicz,M. (1990) The *E.coli dnaK* gene product, the Hsp70 homolog, can reactivate heat-inactivated RNA polymerase in an ATP hydrolysis-dependent manner. *Cell*, **62**, 939–944.

Squires,C.L., Pedersen,S., Ross,B.M. and Squires,C. (1991) ClpB is the *Escherichia coli* heat shock protein F84.1. *J. Bacteriol.*, **173**, 4254–4262.

Tatsuta,T., Tomoyasu,T., Bukau,B., Kitagawa,M., Mori,H., Karata,K. and Ogura,T. (1998) Heat shock regulation in the *ftsH* null mutant of *Escherichia coli*: dissection of stability and activity control mechanisms of σ^{32} *in vivo*. *Mol. Microbiol.*, **30**, 583–593.

Teter,S.A., Houry,W.A., Ang,D., Tradler,T., Rockabrand,D., Fischer,G., Blum,P., Georgopoulos,C. and Hartl,F.U. (1999) Polypeptide flux through bacterial Hsp70: DnaK cooperates with Trigger Factor in chaperoning nascent chains. *Cell*, **97**, 755–765.

Thomas,J.G. and Baneyx,F. (1998) Roles of the *Escherichia coli* small heat shock proteins IbpA and IbpB in thermal stress management: comparison with ClpA, ClpB and HtpG *in vivo*. *J. Bacteriol.*, **180**, 5165–5172.

Tomoyasu,T., Ogura,T., Tatsuta,T. and Bukau,B. (1998) Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in *E.coli*. *Mol. Microbiol.*, **30**, 567–581.

Veinger,L., Diamant,S., Buchner,J. and Goloubinoff,P. (1998) The small heat-shock protein IbpB from *Escherichia coli* stabilizes stress-denatured proteins for subsequent refolding by a multichaperone network. *J. Biol. Chem.*, **273**, 11032–11037.

Zeilstra-Ryalls,J., Fayet,O., Baird,L. and Georgopoulos,C. (1993) Sequence analysis and phenotypic characterization of *groEL* mutations that block λ and T4 bacteriophage growth. *J. Bacteriol.*, **175**, 1134–1143.

Zhu,X., Zhao,X., Burkholder,W.F., Gragerov,A., Ogata,C.M., Gottesman,M. and Hendrickson,W.A. (1996) Structural analysis of substrate binding by the molecular chaperone DnaK. *Science*, **272**, 1606–1614.

Ziemienowicz,A., Skowyra,D., Zeilstra-Ryalls,J., Fayet,O., Georgopoulos, C. and Zylicz,M. (1993) Both the *Escherichia coli* chaperone systems GroEL/Gro/ES and DnaK/DnaJ/GrpE, can activate heat treated RNA polymerase. Different mechanisms for the same activity. *J. Biol. Chem.*, **268**, 25425–25431.

Received July 22, 1999; revised and accepted October 22, 1999

EXHIBIT 2

Heat shock genes – integrating cell survival and death

RICHA ARYA, MOUSHAMI MALLIK and SUBHASH C LAKHOTIA*

Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi 221005, India

**Corresponding author (Fax, 91 542 2368457; Email, lakhotia@bhu.ac.in)*

Heat shock induced gene expression and other cellular responses help limit the damage caused by stress and thus facilitate cellular recovery. Cellular damage also triggers apoptotic cell death through several pathways. This paper briefly reviews interactions of the major heat shock proteins with components of the apoptotic pathways. Hsp90, which acts as a chaperone for unstable signal transducers to keep them poised for activation, interacts with RIP and Akt and promotes NF-κB mediated inhibition of apoptosis; in addition it also blocks some steps in the apoptotic pathways. Hsp70 is mostly anti-apoptotic and acts at several levels like inhibition of translocation of Bax into mitochondria, release of cytochrome c from mitochondria, formation of apoptosome and inhibition of activation of initiator caspases. Hsp70 also modulates JNK, NF-κB and Akt signaling pathways in the apoptotic cascade. In contrast, Hsp60 has both anti- and pro-apoptotic roles. Cytosolic Hsp60 prevents translocation of the pro-apoptotic protein Bax into mitochondria and thus promotes cell survival but it also promotes maturation of procaspase-3, essential for caspase mediated cell death. Our recent *in vivo* studies show that RNAi for the Hsp60D in *Drosophila melanogaster* prevents induced apoptosis. Hsp27 exerts its anti-apoptotic influence by inhibiting cytochrome c and TNF-mediated cell death. $\alpha\beta$ crystallin suppresses caspase-8 and cytochrome c mediated activation of caspase-3. Studies in our laboratory also reveal that absence or reduced levels of the developmentally active as well as stress induced non-coding hsrω transcripts, which are known to sequester diverse hnRNPs and related nuclear RNA-binding proteins, block induced apoptosis in *Drosophila*. Modulation of the apoptotic pathways by Hsps reflects their roles as "weak links" between various "hubs" in cellular networks. On the other hand, non-coding RNAs, by virtue of their potential to bind with multiple proteins, can act as "hubs" in these networks. In view of the integrative nature of living systems, it is not surprising that stress-induced genes, generally believed to primarily function in cell survival pathways, inhibit or even promote cell death pathways at multiple levels to ensure homeostasis at cell and/or organism level. The heat shock genes obviously do much more than merely help cells survive stress.

[Arya R, Mallik M and Lakhotia S C 2007 Heat shock genes – integrating cell survival and death; *J. Biosci.* **32** 595–610]

1. Introduction

As a consequence of being alive, cells of all organisms continuously suffer a variety of "damages" from internal as well as external physico-chemical and biotic factors. Therefore, living systems have evolved a variety of strategies to repair the damage and/or eliminate the damaged components. Heat shock or stress response is a cellular adaptive response, which helps maintain cellular homeostasis under stress. Among the many changes in cellular activity and physiology, the most remarkable event in stressed cells is the production of a highly conserved set of proteins, the Heat Shock or Stress Proteins (Hsps) (Schlesinger *et al* 1982) and certain non-coding RNAs, like the hsrω transcripts in *Drosophila* and the satellite III transcripts in humans (Lakhotia 2003; Jolly and Lakhotia 2006).

The Hsps are broadly classified, on the basis of their apparent molecular weights, amino acid sequences and functions (Nover 1984) into distinct families, viz., Hsp100, Hsp90, Hsp70, Hsp60, Hsp40, small Hsps (sHsp) and Hsp10. Many members of these Hsp families are present

Keywords. Caspases; Hsp90; Hsp70; Hsp60; Hsp27; *hsr-omega*

constitutively (heat shock cognates) in cells while some are expressed only after stress. The induced and the constitutively expressed members of Hsp families are well known as molecular chaperones which i) help in normal folding of various polypeptides, ii) assist mis-folded proteins to attain or regain their native states, iii) regulate protein degradation and/or iv) help in translocation of proteins to different cellular compartments (reviewed in Hartl and Hayer-Hartl 2002).

The above functions imply that the Hsps interact with a very large variety of cellular proteins and thus are important components of cellular networks (Csermely 2004). This is also reflected by their roles, especially of Hsp90, in evolvability (Rutherford and Lindquist 1998). Another interesting example of the integrative roles of Hsps is their intervention in the apoptotic pathways. Apoptosis, one of the programmed cell death pathways, is a natural and essential developmental process, which eliminates redundant or superfluous cells to allow normal patterning (reviewed in Jacobson *et al* 1997). Stressed and damaged cells, if irreparable, also utilize this route to die. Many studies in recent years have shown that the heat shock proteins play critical roles in modulating the apoptotic cascades (Samali and Orrenius 1998; Garrido *et al* 2001; Sreedhar and Csermely 2004; Beere 2005; Kim *et al* 2006a; Didelot *et al* 2006). With a view to understand the wider roles of heat shock genes in cell regulatory pathways, in the following we briefly review how the four major classes of Hsps, viz., Hsp90, Hsp70, Hsp60 and small Hsps modulate apoptotic pathways. We also consider our own recent results relating to roles of one of the Hsp60 forms and the developmentally active and stress-inducible non-coding hsrω transcripts in apoptosis in *Drosophila*. Interlinking of the heat shock response, which primarily repairs the damage, and apoptosis, which eliminates the damaged cells, illustrates the highly integrated nature of the regulatory pathways in living systems.

2. Pathways of apoptosis

Apoptosis is a genetically regulated process of deliberate cell suicide in multicellular organisms. Characteristic features of apoptosis include nuclear condensation, DNA fragmentation, membrane blebbing and breaking of cytoplasm into apoptotic bodies that are removed by phagocytosis (Kerr *et al* 1972; Wyllie *et al* 1980). Multiple triggers (variety of stresses or developmental cues) provoke a cell to undergo apoptosis via one of the two major pathways, the intrinsic or mitochondrial death pathway and the extrinsic or receptor-mediated cell death pathway. Both of these pathways eventually activate effector caspases to execute cell death (Budihardjo *et al* 1999). A brief account of these pathways is given below (for more details, see Yan and Shi 2005). Figure 1 summarizes the major steps in the diverse apoptotic pathways and their links with cell-survival signals.

The *Intrinsic pathway* involves loss of mitochondrial membrane potential in response to death signals, leading to permeabilization of the outer membrane. This triggers release of pro-death molecules like cytochrome c and Smac/Diablo into the cytoplasm. Cytochrome c binds to the Apoptotic protease activating factor-1 (Apaf-1) helping in its oligomerization and recruitment of procaspase-9 to form a functional apoptosome. At the same time, Smac/Diablo inhibit the Inhibitor of Apoptosis Proteins (IAPs). The apoptosome complex proteolytically processes procaspase-9 to an active form, which ultimately leads to cell death by activating the effector caspase-3 (Budihardjo *et al* 1999; Zou *et al* 1999; Van Grup *et al* 2003; Yan and Shi 2005).

The *extrinsic pathway* (figure 1) transduces death signals through the binding of "extra-cellular death ligands" like TNF-α, Fas ligand [FasL]/Apo1L/CD95L, Trail/Apo2L, Apo3L to their respective cell surface receptors. A homotypic interaction takes place between the death domains (DDs) of Tumor Necrosis Factor Receptor-1 (TNFR-1) and Fas-receptors and their respective adaptor molecules, TRADD and FADD (Locksley *et al* 2001; Screaton and Xu 2000; Yan and Shi 2005). Eventually, a Death Inducing Signal Complex (DISC) is formed that activates procaspase-8, which in turn triggers caspase-3 mediated cell death events. Fas-induced apoptotic pathway can also recruit the adaptor protein DAXX instead of FADD, to activate ASK-1 or Apoptosis-Signal-Regulated Kinase-1 (Chang *et al* 1998), which activates SAPK/JNK and thereby, triggers apoptosis (Yang *et al* 1997).

The extrinsic death signals are linked to the intrinsic pathway through the Bcl-2 family of proteins (Gross *et al* 1999), which includes both pro- (e.g. Bax, Bad, Bak, Bid) and anti-apoptotic (e.g. Bcl-XL) members. It is the balance between the proteins of this family with opposing functions that actually decides the release of cytochrome c and Smac/Diablo from the mitochondria (Green and Reed 1998; Willis *et al* 2003).

Death-inducing signals are tightly coupled with survival signals (figure 1). One of the cascades that promote cell survival utilizes a serine/threonine kinase, Akt, which is activated through phosphoinositide 3 kinase, PI(3)K, by various growth factors (Staal 1987; Beere 2001; Paez and Sellers 2003). Phosphorylation stabilizes Akt (Stokoe *et al* 1997), which besides activating the Nuclear Factor-κB (NF-κB) (*see* figure 1), induces phosphorylation of Bad, which results in the latter's disassociation from Bcl-XL. Phosphorylated Bad is sequestered by the cytosolic 14-3-3 protein; this does not permit its translocation into mitochondria and consequently the downstream apoptotic events are not triggered (Zha *et al* 1996).

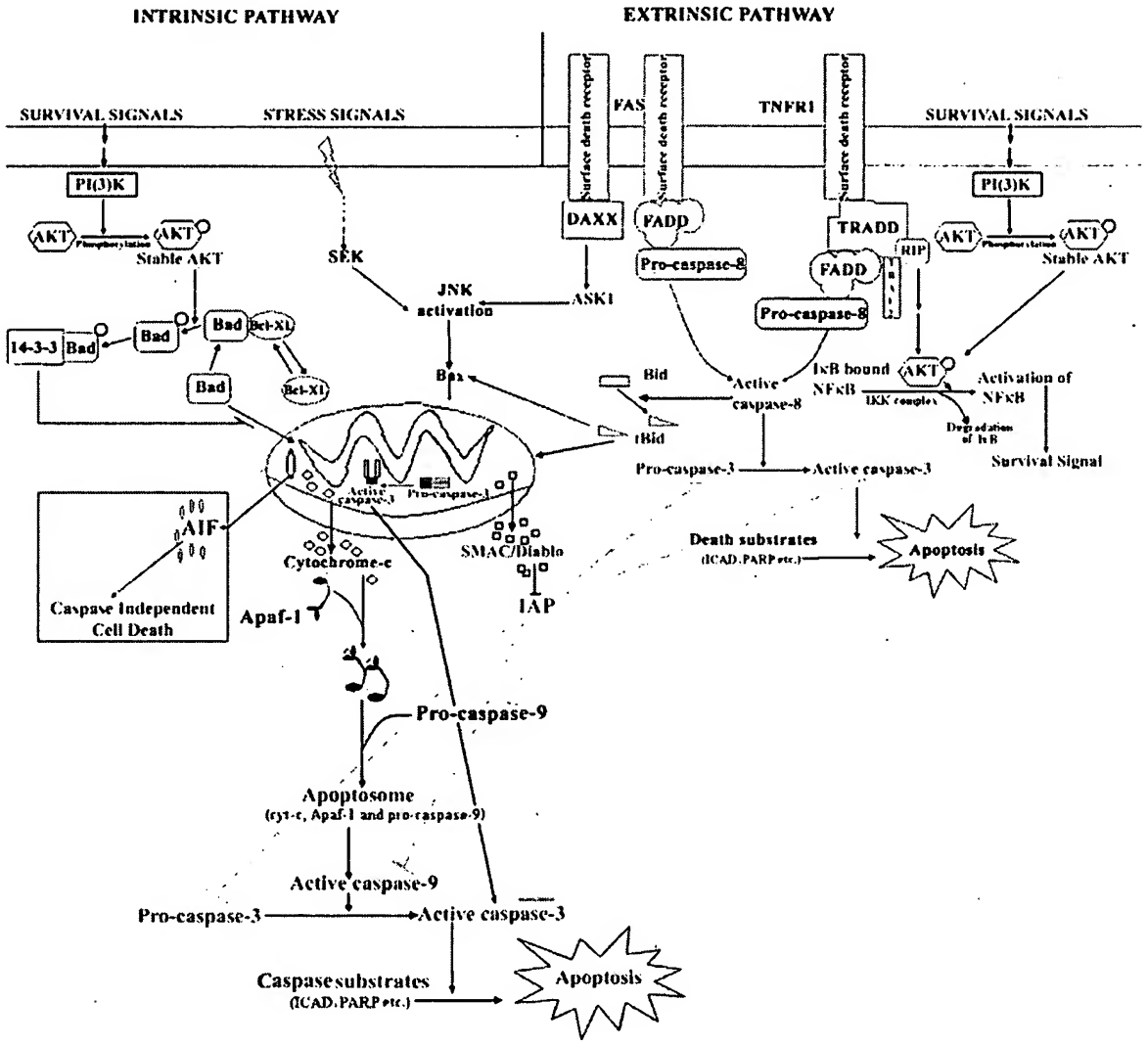


Figure 1. A schematic of intrinsic, extrinsic and AIF-mediated caspase independent pathways of apoptosis and their inter-connections with cell survival pathways. A small blue circle on a box indicates phosphorylated state of the named protein. Arrows indicate promotional while lines with small bar at the end indicate inhibitory actions. The inhibitory effects of IAPs (Inhibitor of Apoptosis Proteins) on pro-caspases and caspases are indicated by dashed lines. See text for details.

Though TNF is a potent inducer of cell death, it may also promote cell survival (figure 1) through NF-κB, an important link in various biological processes including stress response, cell growth or death (Ghosh 1999). Under normal conditions, NF-κB remains bound and sequestered in the cytosol by its inhibitor IκB (DiDonato *et al* 1997; Regnier *et al* 1997). However, upon exposure to stimuli, including TNF, IκB is degraded, resulting in release of NF-κB, which can translocate into nucleus and activate transcription of cell survival genes (Chen *et al* 1996; Wang *et al* 1998). Phosphorylation and thus, inactivation of IκB is mediated by a protein kinase complex, IκB kinase (IKK). TNF triggers activation of IKK through its association with signal transducing molecules like Receptor Interacting Protein, RIP and TRAF2 (Devin *et al* 2000; Zhang *et al* 2000). Activated Akt up-regulates kinase activity of the IKK complex leading to NF-κB mediated cell survival (Ozes *et al* 1999; Romashkova and Makarov 1999).

The core components of cell death pathways are highly conserved among *Caenorhabditis, Drosophila* and humans. *Drosophila* caspases, like those of vertebrates, are broadly grouped into initiator caspases (e.g., Dronc, Dredd and

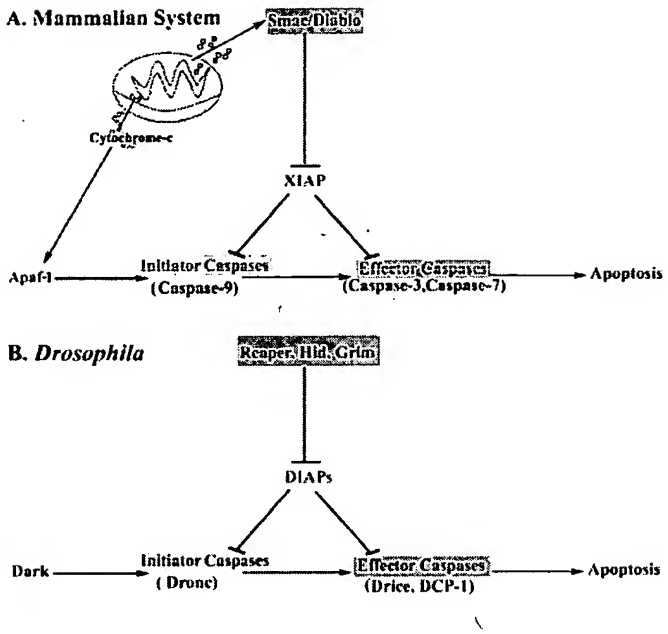


Figure 2. Comparison of the execution steps in mammalian **(A)** and *Drosophila* **(B)** apoptotic machinery. Similar colours indicate functional homology between proteins in mammals and *Drosophila*.

Strica) and effector caspases (e.g. Drice, DCP-1, Decay and Damm) (reviewed in Riedl and Shi 2004; Hay and Guo 2006). In *Drosophila* as well as in vertebrates (figure 2), both groups of caspases are negatively regulated by the IAPs (Deveraux and Reed 1999). Interestingly, in spite of the high conservation of the core cell death machinery, mammals and flies show a remarkable difference in the regulation and execution of apoptosis. In mammalian cells, caspase activation is the primary step of death control, whereas in flies, inactivation of IAPs by upstream pro-apoptotic proteins viz. Reaper, Hid and Grim (functional homologs of vertebrate Smac/Diablo) is the central event. In addition, release of cytochrome c from mitiochondria is required for caspase activation in the vertebrate system whereas role of cytochrome c in *Drosophila* cell death is debated (Kornbluth and White 2005; Hay and Guo 2006).

3. Modulation of apoptosis by different heat shock gene products

On the face of it, Hsps and apoptotic proteins serve two distinct and seemingly opposing functions, viz., survival and death of cells. However, since the cost of survival of a damaged cell or death of a potentially functional cell could be deleterious for the individual, as discussed in the following, the different heat shock proteins and members of the various apoptotic cascades indeed show multi-step networked interactions to fine tune a cell's survival or death.

3.1 *Hsp90*

Hsp90 displays a chaperoning function for unstable signal transducers to keep them poised for activation, although it is not required for their maturation or maintenance (Pratt 1998). In relation to apoptosis, Hsp90 mostly promotes cell survival through its involvement at different steps in the formation of active NF-κB (figure 3). Hsp90 is essential for the stability of RIP, which is recruited by activated TNFR-1 following binding with its ligand, TNF, for sustained NF-κB activity (Lewis *et al* 2000; Chen *et al* 2002). Hsp90 also directly interacts and maintains the activity of Akt by inhibiting its dephosphorylation (Sato *et al* 2000; Basso *et al* 2002). In addition, Hsp90 and its co-chaperone Cdc37 help in the formation of active IKK or Akt complexes, each of which can phosphorylate IκB and thereby cause disassociation of NF-κB from its inhibitor (Chen *et al* 2002).

Apart from its direct role in promoting cell survival pathways, the Hsp90-Akt complex can also indirectly promote cell survival by inhibition of JNK-mediated cell death through phosphorylation and conseuqent inactivation of ASK-1 (figure 3), which is one of the activators of JNK (Zhang *et al* 2005).

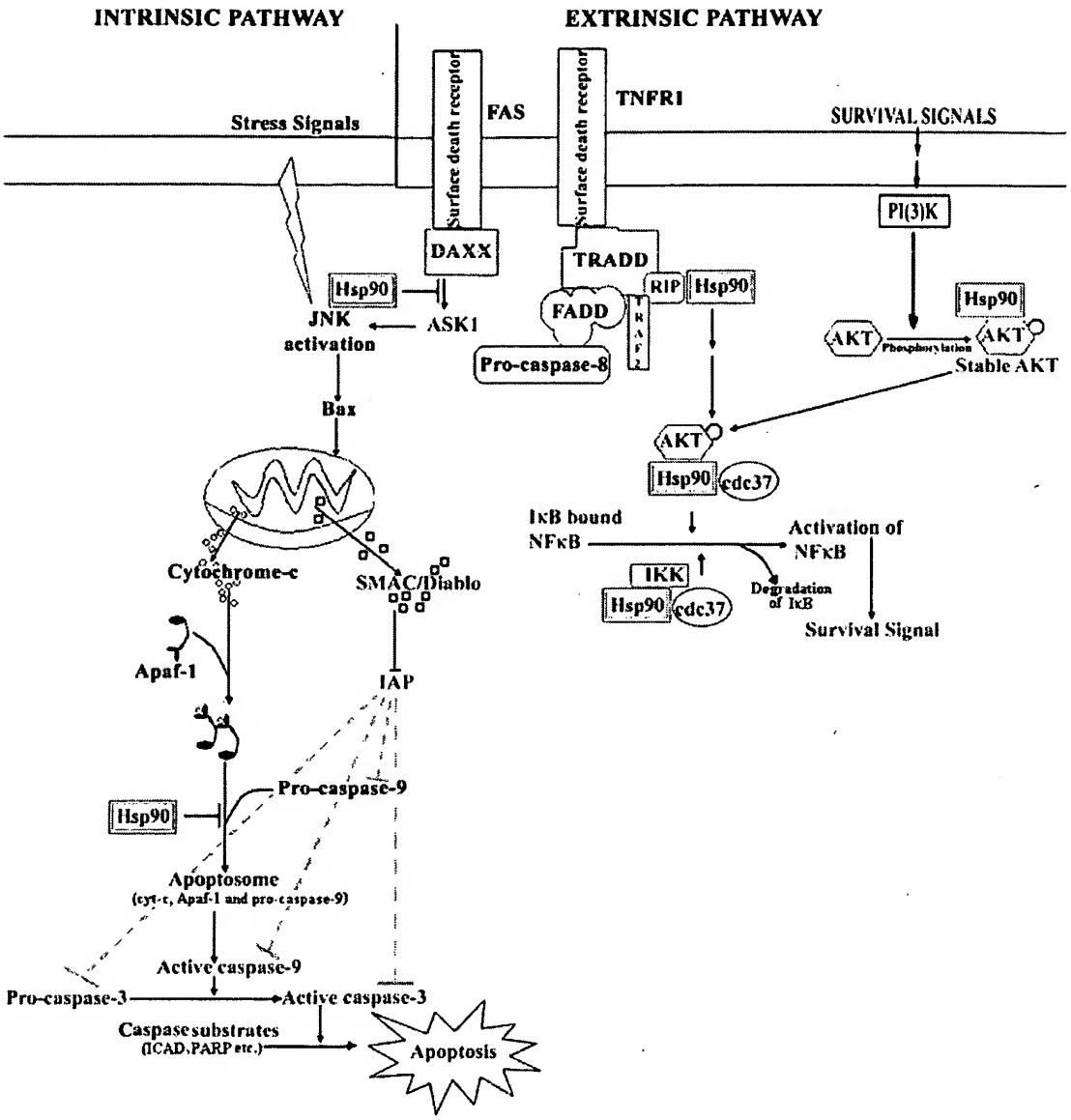


Figure 3. Anti-apoptotic roles of Hsp90 (yellow boxes) are mediated through inhibition of (red lines) activation of ASK-1 or procaspase-9. In addition, Hsp90 (blue boxes) is also an integral component of many complexes that help in cell survival. A small circle on a box indicates phosphorylated state of the given protein. Only a part of the apoptotic and cell survival network (figure 1), relevant to Hsp90 interactions, is shown.

Hsp90 also has a role in modulating the intrinsic pathway of apoptosis (figure 3). It prevents the formation of an active apoptosome complex by inhibiting oligomerization of Apaf-1 (Pandey *et al* 2000a).

Contrary to these mostly cell survival roles of Hsp90, Kim *et al* (2006b) have shown that a direct physical interaction of Hsp90 with the hypoxia-responsive pro-apoptotic protein (HGTD-P) is essential for its translocation into mitochondria for induction of the mitochondrial death pathway.

3.2 *Hsp70*

The Hsp70 family is most diverse and includes many constitutive as well as stress-inducible proteins with

overlapping or unique functions in different cell compartments and in different cellular contexts (reviewed in Fink 1999; Mayer and Bukau 2005). Roles of individual members of the large Hsp70 family in apoptosis are not clearly delineated, since most of the experimental studies on apoptosis have examined the heat shock inducible Hsp70.

Like Hsp90, Hsp70 is also mostly anti-apoptotic. It interacts with the intrinsic and extrinsic pathways of apoptosis at a number of steps and inhibits cell death through chaperone dependent as well as independent activities (figure 4). As an anti-apoptotic protein, Hsp70 protects cells from cytotoxicity induced by TNF, monocytes, oxidative stress, chemotherapeutic agents, ceramide and radiation (Jaattela *et al* 1992; Jaattela and Wissing 1993; Simon *et al* 1995; Karlseder *et al* 1996; Mosser *et al* 1997). The apoptotic cascade stimulated by nitric oxide and heat stress triggers translocation of Bax from cytoplasm to the mitochondria, which is inhibited by over-expression of Hsp70 (Gotoh

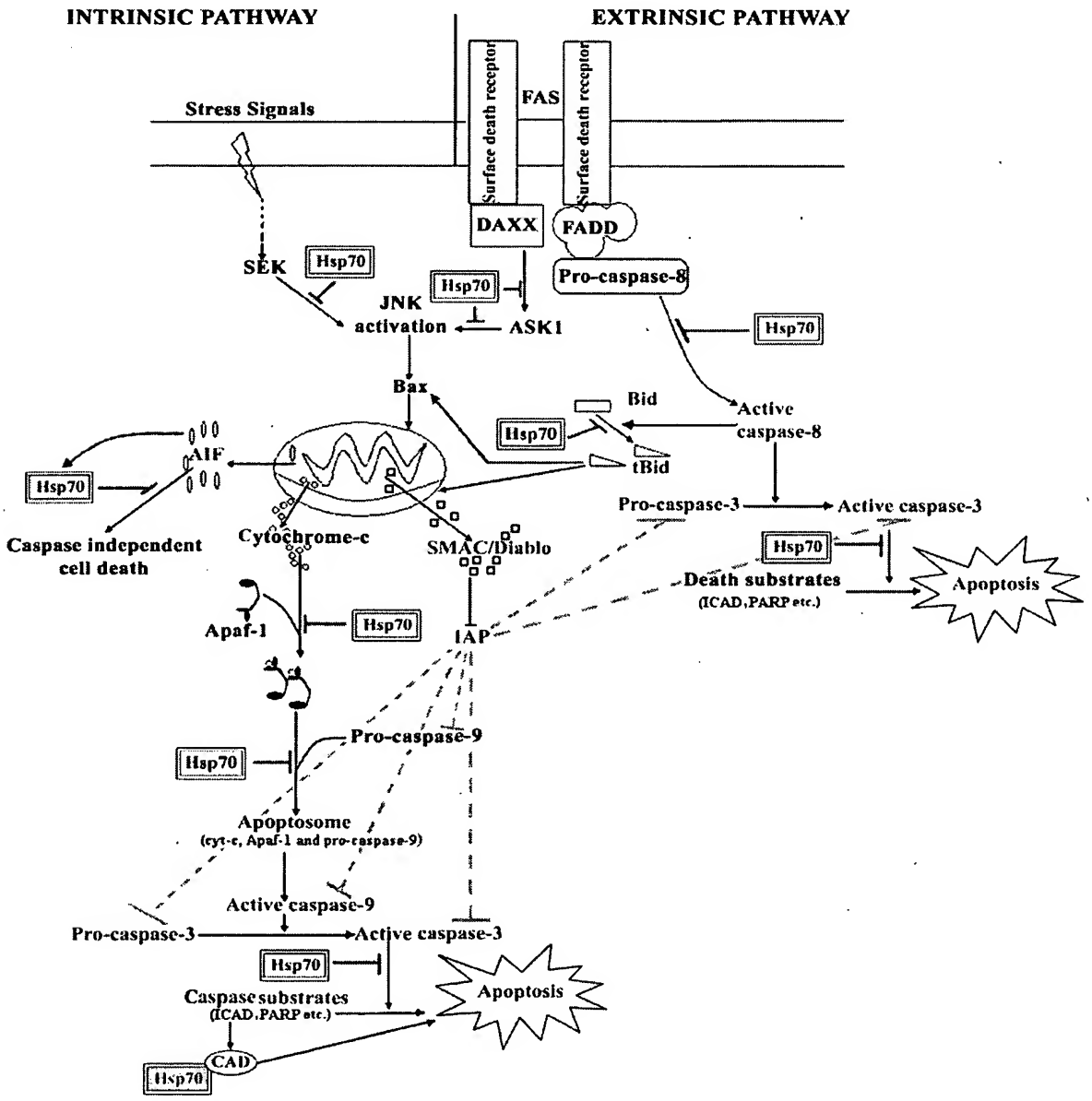


Figure 4. Hsp70 (yellow boxes) is generally anti-apoptotic since it inhibits (red lines) many steps in the apoptotic pathways. However, Hsp70 (green box) has a pro-apoptotic role also since it is required for activity of the caspase-activated DNase (CAD).

et al 2004; Stankiewicz *et al* 2005). Further downstream in the intrinsic pathway, Hsp70 inhibits formation of a functional apoptosome complex by direct interaction with Apaf-1 (Beere *et al* 2000; Saleh *et al* 2000). Hsp70 prevents late caspase dependent events such as activation of cytosolic phospholipase A2 and changes in nuclear morphology; it can also protect cells from forced expression of caspase-3 (Jaattela *et al* 1998).

Hsp70 can, independent of its chaperoning activity, inhibit JNK mediated cell death, by suppressing JNK

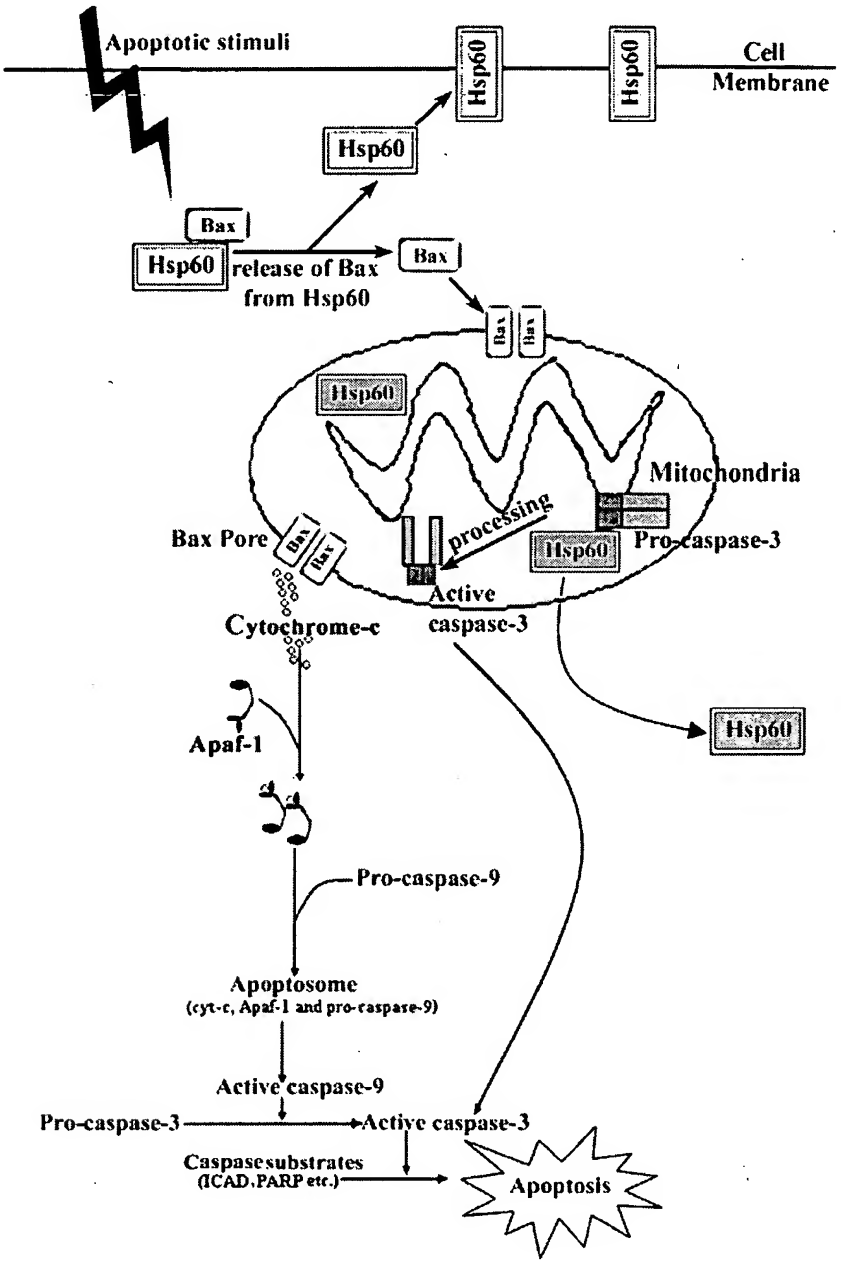


Figure 5. Anti- and pro-apoptotic roles of Hsp60. Cytoplasmic Hsp60 (yellow boxes) inhibits cell death by sequestering Bax and thereby, preventing its translocation to the mitochondrial membrane. On the other hand, mitochondrial Hsp60 (green boxes) promotes maturation of procaspase-3 and thus, has pro-apoptotic role.

phosphorylation either directly and/or through the upstream SEK kinase (Mosser *et al* 1997, 2000; Meriin *et al* 1999; Volloch *et al* 1999). An earlier study in our laboratory (Srivastava 2004) showed that developing eye discs of *Drosophila* expressing a dominant negative mutant form of the major constitutively expressed Hsc70.4 protein displayed high incidence of apoptosis and this was presumably mediated via the JNK pathway.

The caspase-8 mediated activation of Bid (truncated or tBid), which allows Bax to migrate onto the mitochondrial membrane to trigger release of various death factors, is a link between the extrinsic and intrinsic pathways (Li *et al* 1998; Luo *et al* 1998). Hsp70 regulates this activation of Bid independent of its chaperoning function (Gabai *et al* 2002) and thus, can influence both the pathways. In addition, various death-inducing stimuli, viz. TNF-α, Fas and many others are known to cause apoptosis via ASK-1 activation. Hsp70 hampers TNF mediated apoptosis by inhibition of ASK-1 (Park *et al* 2002).

Hsp70 plays anti-apoptotic roles in caspase independent pathway as well through its binding with the Apoptosis Inducing Factor (AIF) released from the mitochondria following death-inducing stimuli and thereby, restricting its translocation into the nucleus (Ravagnan *et al* 2001; Gurbuxani *et al* 2003). In addition, Hsp70 also impedes a lysosome mediated caspase independent cell death pathway since it maintains integrity of the lysosomal membrane and thus, prevents release of cathepsin into the cytosol (Nylandsted *et al* 2004).

In contrast to the above death inhibitory roles, a pro-death role, downstream of caspase-3, has also been ascribed to Hsp70 (figure 3). In Jurkat T cells, Hsp70 enhances TCR/CD3 and Fas/Apo-1/CD95 mediated apoptosis presumably by direct interaction with caspase activated DNase (Liossis *et al* 1997).

3.4 *Hsp60*

The cytosolic and organellar forms of Hsp60 have anti- as well as pro-apoptotic roles (Sarkar *et al* 2006). Hsp60 and its co-chaperonin Hsp10 help prokaryotes survive severe stress (Volker *et al* 1992; Lund 1995). Both these chaperonins, in concert, prevent cell death in neonatal cardiac myocytes by maintaining the mitochondrial membrane integrity and function (Lin *et al* 2001). The cytosolic Hsp60 is mainly anti-apoptotic as it binds with pro-apoptotic Bax and Bak proteins in cardiac myocytes of rat and thus prevents triggering of the apoptotic machinery (Kirchhoff *et al* 2002). Shan *et al* (2003) found Bcl-XL also to be associated with Hsp60 in normal heart tissues. *In vitro* studies demonstrate that Hsp60 provides differential protection against intracellular beta-amyloid induced neuronal dysfunction and cell death by maintaining mitochondrial oxidative phosphorylation (Veereshwarayya *et al* 2006).

Binding of mitochondrial Hsp60 to procaspase-3 enhances its protease-sensitivity and thus makes it more susceptible to the action of cytochrome c and dATP (Samali *et al* 1999; Xanthoudakis *et al* 1999). In agreement with these observations, it was found that patients with esophageal squamous cell carcinomas with high Hsp60 showed higher apoptotic index and thus, enhanced survival (Faried *et al* 2004).

Recent studies in our laboratory suggest that one of the Hsp60 in *Drosophila melanogaster* is necessary for induced apoptosis. The Berkeley *Drosophila* Genome Project has revealed four *Hsp60* genes in *D. melanogaster*, which have been named as *Hsp60A*, *Hsp60B*, *Hsp60C* and *Hsp60D*, respectively (Sarkar and Lakhotia 2005). The *CG16954* or *Hsp60D* gene appears to be necessary for apoptosis. We have generated transgenic flies to conditionally either enhance (*UAS-Hsp60D.WT*) or ablate (*UAS-Hsp60D-RNAi*) the levels of Hsp60D under the control of a desired *Gal4* driver (Brand and Perrimon 1993, Giordano *et al* 2002). It is known that *GMR-Gal4* homozygous flies by themselves show highly disorganized ommatidial arrays in adult flies and a high level of apoptosis in the third instar eye discs (Kramer and Staveley 2003). Our studies (Arya R and Lakhotia S C, unpublished) show that *GMR-Gal4* homozygous flies co-expressing the *Hsp60D–RNAi* transgene, and thus with significantly reduced levels of Hsp60D in developing eye disc cells, have normal ommatidial arrays in adult eyes. Further evidence for essential requirement of Hsp60D in apoptosis is provided by our other (Arya R and Lakhotia S C, unpublished) observation that reduction or loss of Hsp60D protein prevents cell death (*see* figure 6) caused by directed over-expression of some of the key cell death regulators like Reaper, Hid or Grim (Grether *et al* 1995; White *et al* 1996; Hawkins *et al* 2000) in *Drosophila*. Thus it appears that availability of Hsp60D in the cell is essential for execution of induced apoptosis.

3.5 *Hsp27*

Hsp27 belongs to the family of small stress proteins that are constitutively abundant and ubiquitously present. Hsp27 regulates apoptosis through its ability to interact with key components of the apoptotic-signaling pathways (reviewed in Concannon *et al* 2003), particularly those involved in caspase activation (figure 7). Changes in the intracellular redox balance and production of reactive oxygen species initiate the apoptotic cascade through changes in the mitochondria and release of pro-apoptotic factors. Hsp27 can maintain both the redox homeostasis and mitochondrial stability in the cell. Increased expression of Hsp27 during stress response correlates with better survival from cytotoxic stress. It negatively regulates the activation of procaspase-9

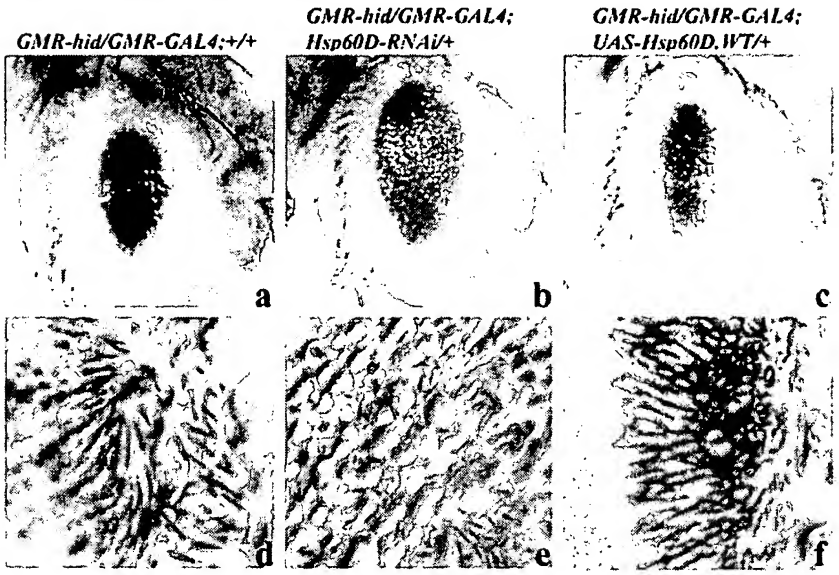


Figure 6. Hsp60D is required for cell death caused by over-expression of Hid in developing *Drosophila* eye discs. Images in **a-c** are photomicrographs while those in **d-f** are nail polish imprints (Arya and Lakhotia 2006) of adult eyes. Over-expression of Hid under *GMR* promoter in developing eye discs results in severe degeneration and reduction in size of adult eyes due to apoptosis of many of the Hid-expressing cells (**a, d**). However, *GMR* driven co-expression of *Hid* and *Hsp60D-RNAi* transgenes and consequent down-regulation of Hsp60D results in partial suppression of Hid mediated eye ablation as evidenced by the enlarged size of eyes (**b**) and formation of some ommatidial arrays (**e**). On the other hand, *GMR* driven over-expression of *HSP60D* in Hid expressing eye discs (**c, f**) resulted in smaller and more damaged eyes.

by sequestering cytosolic cytochrome c from Apaf-1, after its release from mitochondria and thus, prevents assembly of the apoptosome (Bruey *et al* 2000; Concannon *et al* 2001). Hsp27 can block release of cytochrome c from mitochondria in cells exposed to staurosporine, etoposide or cytochalasin D (Paul *et al* 2002). It also mediates inhibition of procaspase-3 activation, most likely through its ability to prevent initiator caspases like caspase-9 from gaining access to the residues whose cleavage is essential for procaspase-3 activation (Pandey *et al* 2000b). In addition, Hsp27 maintains the actin network integrity and hence, prevents translocation of pro-apoptotic factors like activated Bid (tBid) onto the mitochondrial membrane (Paul *et al* 2002).

Hsp27 is reported to block DAXX-mediated apoptosis by preventing its translocation to the membrane and thus, inhibiting its interaction with Fas and ASK-1 (Charette *et al* 2000). Rane *et al* (2003) have suggested that Hsp27 regulates apoptosis of neutrophils through interaction with Akt (Protein Kinase B): Hsp27 is phosphorylated by Akt which results in dissociation of Hsp27 and stabilization of Akt. Disruption of interaction between Akt and Hsp27 impairs Akt activation, which leads to enhanced constitutive apoptosis of neutrophils.

The only known pro-apoptotic role of Hsp27 so far is that it enhances TNF-induced apoptosis by inhibiting IκB degradation and thereby, preventing NF-κB mediated cell survival (Kammanadiminti and Chadee 2006).

αβ-crystallin, a small Hsp family member closely related to Hsp27 (Ingolia and Craig 1982), is constitutively expressed in many tissues and is especially abundant in eye lens, heart and in muscles. αβ-crystallin interferes in the processing of the precursor of procaspase-3 (Mao *et al* 2001; Kamradt *et al* 2001; Alge *et al* 2002). Additionally, αβ-crystallin inhibits apoptosis through sequestration of Bax and Bcl-Xs in the cytoplasm (Mao *et al* 2004). Over expression of αβ-crystallin can also inhibit apoptosis caused by RAS activation (Li *et al* 2005).

3.6 *Non-coding hsrω transcripts in* Drosophila

The *hsrω* gene of *D. melanogaster* is developmentally active in most cell types and is also one of the most strongly induced genes following heat shock. This gene produces several non-coding RNAs as functional end products (reviewed in Lakhotia 2003). The >10 Kb nucleus limited hsrω-n transcript of this gene is dynamically associated with several different hnRNPs in the nucleus to form fine nucleoplasmic omega speckles and it has been suggested that the omega speckles regulate nuclear trafficking and availability of hnRNPs and other related RNA binding

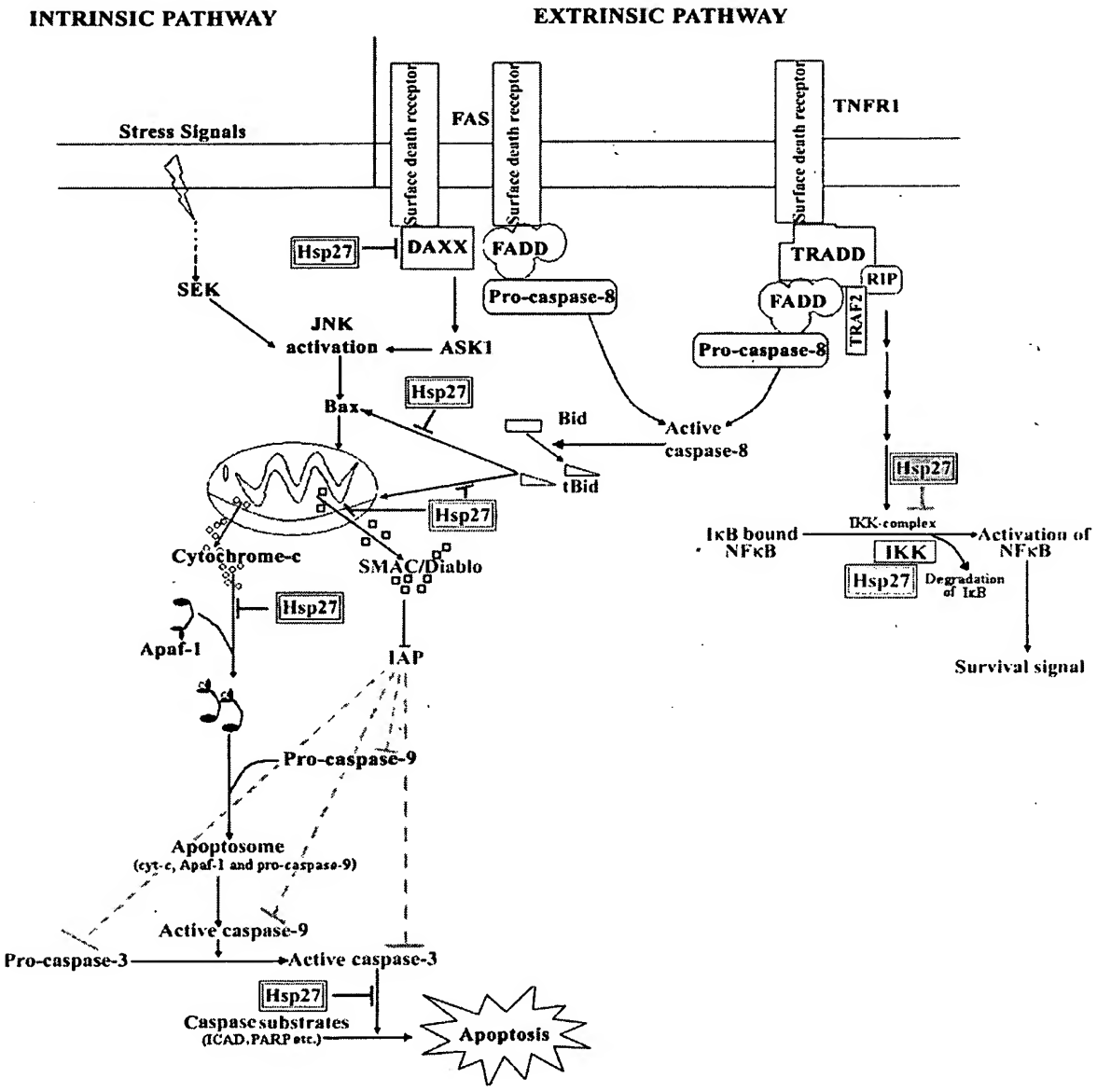


Figure 7. Anti-apoptotic (yellow) and pro-apoptotic (green) roles of Hsp27 in intrinsic and extrinsic cell death pathways (see text for details).

proteins in the cell (Lakhotia *et al* 1999; Prasanth *et al* 2000).

It is known that expression of expanded polyglutamine (polyQ) containing proteins is cytotoxic to neuronal cells, which ultimately die because of apoptosis (Warrick *et al* 1998; Faber *et al* 1999; Paulson *et al* 2000). Recent studies in our laboratory demonstrated that mis- or over-expression of the non-coding hsrω transcripts enhances polyQ-induced neurodegeneration (Sengupta and Lakhotia 2006). To gain a better understanding of the physiological roles of the non-coding *hsrω* gene, we have employed GAL4 driven over-expression (EP mediated) or ablation (RNAi mediated) of hsrω transcripts and examined the ensuing developmental/ phenotypic abnormalities. Our results (Mallik M and Lakhotia S C, in preparation) show that ablation of the hsrω-n transcripts through RNAi in developing eye disc

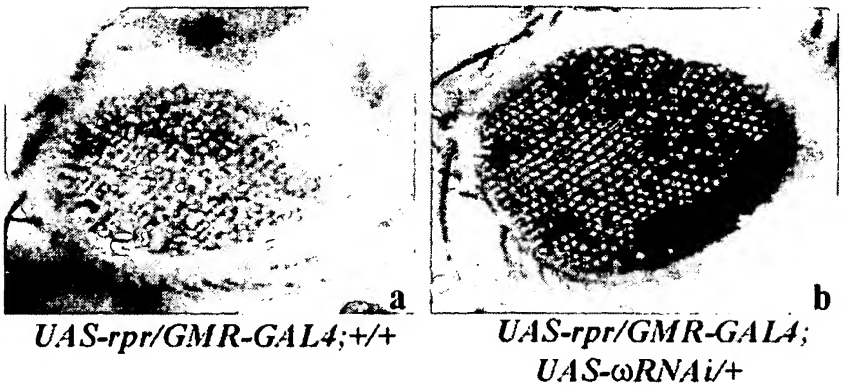


Figure 8. Absence of the non-coding hsrω transcripts suppresses induced cell death. Ectopic expression of the apical cell death trigger Reaper in photoreceptor neurons and accessory pigment cells of developing eye discs under control of the *GMR-GAL4* driver results in massive cell death in the developing retina and causes rough and reduced eyes in adult flies (**a**). However, reduction in the levels of the omega-n transcripts through RNAi in eye disc cells expressing Reaper restores the adult eye morphology to near wild type (**b**).

cells expressing expanded polyQ substantially rescues neurodegeneration in eyes. In order to see if this rescue is due to inhibition of apoptosis, we have examined effects of alterations in the levels of the hsrω transcripts on apoptosis following activation of Reaper and other apical cell death triggers in developing retina. It is seen (Mallik M and Lakhotia S C, in preparation) that, as expected, *GMR-Gal4* driven over-expression of Reaper in developing wild type eye discs has severe effects on morphology and pigmentation of adult eyes (figure 8a); however, RNAi of hsrω transcripts in eye discs over-expressing Reaper results in near complete absence of eye defects (figure 8b). We have observed similar inhibitory effect of RNAi of the non-coding hsrω transcripts on apoptosis triggered by other cell death stimuli (Mallik M and Lakhotia S C, in preparation). This novel involvement of a non-coding RNA in apoptosis opens new paradigm/s for integration of various cellular activities.

4. Epilogue

The above brief survey clearly shows that many of the heat shock proteins (constitutive as well as stress induced) interact with members of the apoptotic cascades to inhibit or promote cell death. It is significant that members of each of the Hsp families not only act at multiple levels but even in opposing directions.

At any given time, a living cell is a multi-tasking system in which multitudes of external and internal information/ signals are processed in parallel but the responses to any one or more of these activities have to be integrative. Hsps by the nature of their functions, appear to be significant integrators or co-ordinators of cellular signals and responses. They provide "weak" but very important links between many "hubs" and thus, help transmission of information/signal across different paths in cellular networks (Csermely 2004; Korcsmáros *et al* 2007).

The intricately networked organization of the various functional and structural compartments of a cell implies that damage or unfavourable perturbations in any one compartment can have varying consequences for other compartments as well. In view of the existence of multiple signals that can activate one or the other apoptotic pathways, any of these compartments can trigger cell death depending upon the extent of the unfavourable signals received. Excessive cell death can be very "costly" for the organism and therefore, appropriate checks need to be exercised before the apoptotic cascade is permitted to execute cell death. Hsps provide one of the check systems. The significance of the modulatory roles of the various Hsps at multiple steps in apoptotic cascades lies in the fact that as weak "links" (Csermely 2004), they can tilt the balance in subtle but precise ways in favour of cell survival or cell death in a context-dependent manner.

In addition to the heat shock induced genes, the Heat shock transcription factors (HSFs), which primarily regulate the activation of heat shock genes following stress (Pirkkala *et al* 2001), are also now known to modulate the apoptotic cascade. Xia *et al* (2000) have shown that in HeLa cell line the activated Fas death receptor transactivates HSF1 which induces heat shock genes. Interestingly, despite the presence of activated Fas, these cells are insensitive to Fas-mediated cell death but over-expression of constitutively active HSF1 sensitizes them for cell death; this suggests novel interactions between the stress protein mediated cell survival and activated Fas-dependent cell killing (Xia *et al* 2000). Boellman *et al* (2004) demonstrated that the DAXX protein,

a modulator of apoptosis, also induces the transactivation of HSF1 under stress conditions in human cell lines. Recently HSF1 mediated cell death has been described in male germ cells (Hayashida *et al* 2006, Vydra *et al* 2006). Thus HSFs emerge as novel modulators of cell survival or cell death.

It is becoming increasingly appreciated that non-coding RNA species have important regulatory functions in eukaryotic cells (Lakhotia 1996, 2003; Eddy 2001; Mattick 2004; Prasanth and Spector 2007). Recent observations in our laboratory (Mallik M and Lakhotia S C, in preparation) that a large non-coding RNA, expressed constitutively and further activated by stress, can also modulate induced apoptosis add a new dimension to the regulation of cell death and survival. The hsrω-n transcripts are known to bind with a variety of hnRNPs, several other RNA-processing/binding proteins and Hsp90 (Lakhotia *et al* 1999; Prasanth *et al* 2000; Lakhotia 2003; Jolly and Lakhotia 2006) and thus, can potentially affect all those cellular activities in which these proteins are directly or indirectly involved. hnRNPs are known to modulate apoptosis (Jiang *et al* 1998; Charroux *et al* 1999; Shchors *et al* 2002; Hermann *et al* 2001; Schwerk and Schulze-Osthoff 2005) mostly through regulation of alternative splicing. It is likely that down-regulation of hsrω-n transcripts by RNAi, affects the nuclear hnRNP metabolism and this may affect apoptosis. Association of hsrω transcripts with Hsp90 may also be significant in this context. Association or a direct interaction of the hsrω-n transcripts with any of the caspases or the main components of the apoptotic pathways is not yet known, although our observations (Mallik M and Lakhotia S C, in preparation) indicate that reduced levels of hsrω-n transcripts prevent activation of JNK pathway and this in turn may block apoptosis. Irrespective of the mode of action, it is significant that a non-coding RNA affects a vital process. Since binding of proteins to target nucleic acids is dependent upon short motifs, we believe that RNA molecules can have multiple protein-binding potential and thus, non-coding RNA species like the hsrω-n transcripts can function as hubs in the cellular network systems with the various RNA-binding proteins providing the links with other hubs.

The high conservation of Hsp families from bacteria through higher eukaryotes shows that these are very ancient proteins and thus, had many opportunities to evolve interactions with other newly emerging proteins. Their function as molecular chaperones by itself also requires their involvement in a large variety of cellular events (Korcsmáros *et al* 2007). In addition, they participate in many cellular events in a chaperone independent manner as well. It is obvious that contrary to their generic names, the Hsps are not just reversing stress-induced cellular damage or helping nascent polypeptides to fold appropriately, but have more global roles in cell metabolism. Likewise, stress-induced non-coding transcripts also appear to have "conserved functions" (Jolly and Lakhotia 2006). An integrative understanding of the various functions of Hsps and non-coding transcripts will help us know better the networks operating in biological systems.

Acknowledgements

RA is receiving Senior Research Fellowship from the Council of Scientific and Industrial Research, New Delhi. MM is SPM Fellow (Council of Scientific and Industrial Research, New Delhi). Research work in the laboratory is supported by grants from the Department of Biotechnology and Department of Science and Technology, New Delhi to SCL.

References

Alge C S, Priglinger S G, Neubauer A S, Kampik A, Zillig M, Bloemendal H and Welge-Lussen U 2002 Retinal pigment epithelium is protected against apoptosis by alphaB-crystallin; *Invest. Ophthalmol. Vis. Sci.* **43** 3575–3582

Arya R and Lakhotia S C 2006 A simple nail polish imprint technique for examination of external morphology of *Drosophila* eyes; *Curr. Sci.* **90** 1179–1180

Basso A D, Solit D B, Chiosis G, Giri B, Tsichlis P and Rosen N 2002 Akt forms an intracellular complex with heat shock protein 90 (Hsp90) and Cdc37 and is destabilized by inhibitors of Hsp90 function; *J. Biol. Chem.* **277** 39858–39866

Beere H M 2001 Stressed to death: regulation of apoptotic signaling pathways by the heat shock proteins; *Sci. STKE* **2001** RE1

Beere H M 2005 Death versus survival: functional interaction between the apoptotic and stress-inducible heat shock protein pathways; *J. Clin. Invest.* **115** 2633–2639

Beere H M, Wolf B B, Cain K, Mosser D D, Mahboubi A, Kuwana T, Tailor P, Morimoto R I, Cohen G M and Green D R 2000 Heat-shock protein 70 inhibits apoptosis by preventing recruitment of procaspase-9 to the Apaf-1 apoptosome; *Nat. Cell Biol.* **2** 469–475

Boellmann F, Guettouche T, Guo Y, Fenna M, Mnayer L and Voellmy R 2004 DAXX interacts with heat shock factor 1 during stress activation and enhances its transcriptional activity; *Proc Natl Acad Sci USA* **101** 4100–4105

Brand A H and Perrimon N 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes; *Development* **118** 401–415

Bruey J M, Ducasse C, Bonniaud P, Ravagnan L, Susin S A, Diaz-Latoud C, Gurbuxani S, Arrigo A P, Kroemer G, Solary E and Garrido C 2000 Hsp27 negatively regulates cell death by interacting with cytochrome c; *Nat. Cell Biol.* **2** 645–652

Budihardjo I, Oliver H, Lutter M, Luo X and Wang X 1999 Biochemical pathways of caspase activation during apoptosis; *Annu. Rev. Cell Dev. Biol.* **15** 269–290

Chang H Y, Nishitoh H, Yang X, Ichijo H and Baltimore D 1998 Activation of apoptosis signal-regulating kinase 1 (ASK1) by the adapter protein Daxx; *Science* **281** 1860–1863

Charette S J, Lavoie J N, Lambert H and Landry J 2000 Inhibition of Daxx-mediated apoptosis by heat shock protein 27; *Mol. Cell Biol.* **20** 7602–7612

Charroux B, Angelats C, Fasano L, Kerridge S and Vola C 1999 The levels of the bancal product, a *Drosophila* homologue of vertebrate hnRNP K protein, affect cell proliferation and apoptosis in imaginal disc cells; *Mol. Cell. Biol.* **19** 7846–7856

Chen G, Cao P and Goeddel D V 2002 TNF-induced recruitment and activation of the IKK complex require Cdc37 and Hsp90; *Mol. Cell* **9** 401–410

Chen Z J, Parent L and Maniatis T 1996 Site-specific phosphorylation of IkappaBalpha by a novel ubiquitination-dependent protein kinase activity; *Cell* **84** 853–862

Concannon C G, Gorman A M and Samali A 2003 On the role of Hsp27 in regulating apoptosis; *Apoptosis* **8** 61–70

Concannon C G, Orrenius S and Samali A 2001 Hsp27 inhibits cytochrome c-mediated caspase activation by sequestering both pro-caspase-3 and cytochrome c; *Gene Expr.* **9** 195–201

Csermely P 2004 Strong links are important, but weak links stabilize them; *Trends Biochem. Sci.* **29** 331–334

Deveraux Q L and Reed J C 1999 IAP family proteins--suppressors of apoptosis; *Genes Dev.* **13** 239–252

Devin A, Cook A, Lin Y, Rodriguez Y, Kelliher M and Liu Z 2000 The distinct roles of TRAF2 and RIP in IKK activation by TNF-R1: TRAF2 recruits IKK to TNF-R1 while RIP mediates IKK activation; *Immunity* **12** 419–429

Didelot C, Schmitt E, Brunet M, Maingret L, Parcellier A and Garrido C 2006 Heat shock proteins: endogenous modulators of apoptotic cell death; *Handb. Exp. Pharmacol.* 171–198

DiDonato J A, Hayakawa M, Rothwarf D M, Zandi E and Karin M 1997 A cytokine-responsive IkappaB kinase that activates the transcription factor NF-kappaB; *Nature (London)* **388** 548–554

Eddy S R 2001 Non-coding RNA genes and the modern RNA world; *Nat. Rev. Genet.* **2** 919–929

Faber P W, Alter J R, MacDonald M E and Hart A C 1999 Polyglutamine-mediated dysfunction and apoptotic death of a Caenorhabditis elegans sensory neuron; *Proc. Natl. Acad. Sci. USA* **96** 179–184

Faried A, Sohda M, Nakajima M, Miyazaki T, Kato H and Kuwano H 2004 Expression of heat-shock protein Hsp60 correlated with the apoptotic index and patient prognosis in human oesophageal squamous cell carcinoma; *Eur. J. Cancer* **40** 2804–2811

Fink A L 1999 Chaperone-mediated protein folding; *Physiol. Rev.* **79** 425–449

Gabai V L, Mabuchi K, Mosser D D and Sherman M Y 2002 Hsp72 and stress kinase c-jun N-terminal kinase regulate the bid-dependent pathway in tumor necrosis factor-induced apoptosis; *Mol. Cell. Biol.* **22** 3415–3424

Garrido C, Gurbuxani S, Ravagnan L and Kroemer G 2001 Heat shock proteins: endogenous modulators of apoptotic cell death; *Biochem. Biophys. Res. Commun.* **286** 433–442

Ghosh S 1999 Regulation of inducible gene expression by the transcription factor NF-kB. Immunol.; *Immunol. Res.* **19** 183–189

Giordano E, Rendina R, Peluso I and Furia M 2002 RNAi triggered by symmetrically transcribed transgenes in *Drosophila melanogaster*; *Genetics* **160** 637–648

Gotoh T, Terada K, Oyadomari S and Mori M 2004 hsp70-DnaJ chaperone pair prevents nitric oxide- and CHOP-induced apoptosis by inhibiting translocation of Bax to mitochondria; *Cell Death Differ.* **11** 390–402

Green D R and Reed J C 1998 Mitochondria and apoptosis; *Science* **281** 1309–1312

Grether M E, Abrams J M, Agapite J, White K and Steller H 1995 The head involution defective gene of *Drosophila melanogaster* functions in programmed cell death; *Genes Dev.* **9** 1694–1708

Gross A, McDonnell J M and Korsmeyer S J 1999 BCL-2 family members and the mitochondria in apoptosis; *Genes Dev.* **13** 1899–1911

Gurbuxani S, Schmitt E, Cande C, Parcellier A, Hammann A, Daugas E, Kouranti I, Spahr C, Pance A, Kroemer G and Garrido C 2003 Heat shock protein 70 binding inhibits the nuclear import of apoptosis-inducing factor; *Oncogene* **22** 6669–6678

Hartl F U and Hayer-Hartl M 2002 Molecular chaperones in the cytosol: from nascent chain to folded protein; *Science* **295** 1852–1858

Hawkins C J, Yoo S J, Peterson E P, Wang S L, Vernooy S Y and Hay B A 2000 The *Drosophila* caspase DRONC cleaves following glutamate or aspartate and is regulated by DIAP1, HID, and GRIM; *J. Biol. Chem.* **275** 27084–27093

Hay B A and Guo M 2006 Caspase-dependent cell death in *Drosophila*; *Annu. Rev. Cell Dev. Biol.* **22** 623–650

Hayashida N, Inouye S, Fujimoto M, Tanaka Y, Izu H, Takaki E, Ichikawa H, Rho J and Nakai A 2006 A novel HSF1-mediated death pathway that is suppressed by heat shock proteins; *EMBO J.* **25** 4773–4783

Hermann R, Hensel F, Muller E C, Keppler M, Souto-Carneiro M, Brandlein S, Muller-Hermelink H K and Vollmers H P 2001 Deactivation of regulatory proteins hnRNP A1 and A2 during SC-1 induced apoptosis; *Hum. Antibodies* **10** 83–90

Ingolia T D and Craig E A 1982 Four small *Drosophila* heat shock proteins are related to each other and to mammalian alpha-crystallin; *Proc. Natl. Acad. Sci. USA* **79** 2360–2364

Jaattela M and Wissing D 1993 Heat-shock proteins protect cells from monocyte cytotoxicity: possible mechanism of self-protection; *J. Exp. Med.* **177** 231–236

Jaattela M, Wissing D, Bauer P A and Li G C 1992 Major heat shock protein hsp70 protects tumor cells from tumor necrosis factor cytotoxicity; *EMBO J.* **11** 3507–3512

Jaattela M, Wissing D, Kokholm K, Kallunki T and Egeblad M 1998 Hsp70 exerts its anti-apoptotic function downstream of caspase-3-like proteases; *EMBO J.* **17** 6124–6134

Jacobson M D, Weil M and Raff M C 1997 Programmed cell death in animal development; *Cell* **88** 347–354

Jiang Z H, Zhang W J, Rao Y and Wu J Y 1998 Regulation of Ich-1 pre-mRNA alternative splicing and apoptosis by mammalian splicing factors; *Proc. Natl. Acad. Sci. USA* **95** 9155–9160

Jolly C and Lakhotia S C 2006 Human sat III and *Drosophila* hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells; *Nucleic Acids Res.* **34** 5508–5514

Kammanadiminti S J and Chadee K 2006 Suppression of NF-κB activation by Entameba histolytica in intestinal epithelial cells is mediated by heat shock protein 27; *J. Biol. Chem.* **281** 26029–26080

Kamradt M C, Chen F and Cryns V L 2001 The small heat shock protein alpha B-crystallin negatively regulates cytochrome c- and caspase-8-dependent activation of caspase-3 by inhibiting its autoproteolytic maturation; *J. Biol. Chem.* **276** 16059–16063

Karlseder J, Wissing D, Holzer G, Orel L, Sliutz G, Auer H, Jaattela M and Simon M M 1996 HSP70 overexpression mediates the escape of a doxorubicin-induced G2 cell cycle arrest; *Biochem. Biophys. Res. Commun.* **220** 153–159

Kerr J F, Wyllie A H and Currie A R 1972 Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics; *Br. J. Cancer* **26** 239–257

Kim H P, Morse D and Choi A M 2006a Heat-shock proteins: new keys to the development of cytoprotective therapies; *Expert Opin. Ther. Targets* **10** 759–769

Kim J Y, Kim S M, Ko J H, Yim J H and Park J H 2006b Interaction of pro-apoptotic protein HGTD-P with heat shock protein 90 is required for induction of mitochondrial apoptotic cascades; *FEBS Lett.* **580** 3270–3275

Kirchhoff S R, Gupta S and Knowlton A A 2002 Cytosolic heat shock protein 60, apoptosis, and myocardial injury; *Circulation* **105** 2899–2904

Korcsmáros T, Kovács I A , Szalay M S, and Csermely P 2007 Molecular chaperones: The modular evolution of cellular networks.; *J. Biosci.* **32** 441–446

Kornbluth S and White K 2005 Apoptosis in *Drosophila*: neither fish nor fowl (nor man, nor worm); *J. Cell. Sci.* **118** 1779–1787

Kramer J M and Staveley B E 2003 GAL4 causes developmental defects and apoptosis when expressed in the developing eye of *Drosophila melanogaster*; *Genet. Mol. Res.* **2** 43–47

Lakhotia S C 1996 RNA polymerase II dependent genes that do not code for protein; *Indian J. Biochem. Biophys.* **133** 93–102

Lakhotia S C 2003 The noncoding developmentally active and stress-inducible *hsrω* gene of *Drosophila melanogaster* integrates Post-transcriptional processing of other nuclear transcripts; in *Noncoding RNAs: Molecular biology and molecular medicine* (eds) J Barciszewski and VA Erdmann (New York: Kluwer Academic/ Plenum Publishers) pp 209–219

Lakhotia S C, Ray P, Rajendra, T K and Prasanth K V 1999 The non-coding transcripts of hsrω gene in *Drosophila*: Do they regulate trafficking and availability of nuclear RNA processing factors?; *Curr. Sci.* **77** 553–563

Lewis J, Devin A, Miller A, Lin Y, Rodriguez Y, Neckers L and Liu Z G 2000 Disruption of hsp90 function results in degradation of the death domain kinase, receptor-interacting protein (RIP), and blockage of tumor necrosis factor-induced nuclear factor-kappaB activation; *J. Biol. Chem.* **275** 10519–10526

Li D W, Liu J P, Mao Y W, Xiang H, Wang J, Ma W Y, Dong Z, Pike H M, Brown R E and Reed J C 2005 Calcium-activated RAF/MEK/ERK signaling pathway mediates p53-dependent apoptosis and is abrogated by alpha B-crystallin through inhibition of RAS activation; *Mol. Biol. Cell.* **16** 4437–4453

Li H, Zhu H, Xu C J and Yuan J 1998 Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis; *Cell* **94** 491–501

Lin K M, Lin B, Lian I Y, Mestril R, Scheffler I E and Dillmann W H 2001 Combined and individual mitochondrial HSP60 and HSP10 expression in cardiac myocytes protects mitochondrial function and prevents apoptotic cell deaths induced by simulated ischemia-reoxygenation; *Circulation* **103** 1787–1792

Liossis S N, Ding X Z, Kiang J G and Tsokos G C 1997 Overexpression of the heat shock protein 70 enhances the TCR/CD3- and Fas/Apo-1/CD95-mediated apoptotic cell death in Jurkat T cells; *J. Immunol.* **158** 5668–5675

Locksley R M, Killeen N and Lenardo M J 2001 The TNF and TNF receptor superfamilies: integrating mammalian biology; *Cell* **104** 487–501

Lund P A 1995 The roles of molecular chaperones in vivo; *Essays Biochem.* **29** 113-123

Luo X, Budihardjo I, Zou H, Slaughter C and Wang X 1998 Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors; *Cell* **94** 481–490

Mao Y W, Liu J P, Xiang H and Li D W 2004 Human alphaA- and alphaB-crystallins bind to Bax and Bcl-X(S) to sequester their translocation during staurosporine-induced apoptosis; *Cell Death Differ.* **11** 512–526

Mao Y W, Xiang H, Wang J, Korsmeyer S, Reddan J and Li D W 2001 Human bcl-2 gene attenuates the ability of rabbit lens epithelial cells against H2O2-induced apoptosis through down-regulation of the alpha B-crystallin gene; *J. Biol. Chem.* **276** 43435–43445

Mattick J S 2004 RNA regulation: a new genetics? *Nat. Rev. Genet.* **5** 316–323

Mayer M P and Bukau B 2005 Hsp70 chaperones: cellular functions and molecular mechanism; *Cell Mol. Life Sci.* **62** 670–684

Meriin A B, Yaglom J A, Gabai V L, Zon L, Ganiatsas S, Mosser D D and Sherman M Y 1999 Protein-damaging stresses activate c-Jun N-terminal kinase via inhibition of its dephosphorylation: a novel pathway controlled by HSP72; *Mol. Cell Biol.* **19** 2547–2555

Mosser D D, Caron A W, Bourget L, Denis-Larose C and Massie B 1997 Role of the human heat shock protein hsp70 in protection against stress-induced apoptosis; *Mol. Cell Biol.* **17** 5317–5327

Mosser D D, Caron A W, Bourget L, Meriin A B, Sherman M Y, Morimoto R I and Massie B 2000 The chaperone function of hsp70 is required for protection against stress-induced apoptosis; *Mol. Cell Biol.* **20** 7146–7159

Nover L 1984 *Heat shock response of eucaryotic cells* (Berlin: Springer-Verlag)

Nylandsted J, Gyrd-Hansen M, Danielewicz A, Fehrenbacher N, Lademann U, Hoyer-Hansen M, Weber E, Multhoff G, Rohde M and Jaattela M 2004 Heat shock protein 70 promotes cell survival by inhibiting lysosomal membrane permeabilization; *J. Exp. Med.* **200** 425–435

Ozes O N, Mayo L D, Gustin J A, Pfeffer S R, Pfeffer L M and Donner D B 1999 NF-kappaB activation by tumour necrosis factor requires the Akt serine-threonine kinase; *Nature (London)* **401** 82–85

Paez J and Sellers W R 2003 PI3K/PTEN/AKT pathway. A critical mediator of oncogenic signaling; *Cancer Treat. Res.***115** 145–167

Pandey P, Saleh A, Nakazawa A, Kumar S, Srinivasula S M, Kumar V, Weichselbaum R, Nalin C, Alnemri E S, Kufe D

and Kharbanda S 2000a Negative regulation of cytochrome c-mediated oligomerization of Apaf-1 and activation of procaspase-9 by heat shock protein 90; *EMBO J.* **19** 4310–4322

Pandey P, Farber R, Nakazawa A, Kumar S, Bharti A, Nalin C, Weichselbaum R, Kufe D and Kharbanda S 2000b Hsp27 functions as a negative regulator of cytochrome c-dependent activation of procaspase-3; *Oncogene* **19** 1975–1981

Park H S, Cho S G, Kim C K, Hwang H S, Noh K T, Kim M S, Huh S H, Kim M J, Ryoo K, Kim E K, Kang W J, Lee J S, Seo J S, Ko Y G, Kim S and Choi E J 2002 Heat shock protein hsp72 is a negative regulator of apoptosis signal-regulating kinase 1; *Mol. Cell Biol.* **22** 7721–7730

Paul C, Manero F, Gonin S, Kretz-Remy C, Virot S and Arrigo A P 2002 Hsp27 as a negative regulator of cytochrome C release; *Mol. Cell Biol.* **22** 816–834

Paulson H L, Bonini N M and Roth K A 2000 Polyglutamine disease and neuronal cell death; *Proc. Natl. Acad. Sci. USA* **97** 12957–12958

Pirkkala L, Nykanen P and Sistonen L 2001 Roles of the heat shock transcription factors in regulation of the heat shock response and beyond; *Faseb J.* **15** 1118–1131

Prasanth K V and Spector D L 2007 Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum; *Genes Dev.* **21** 11–42

Prasanth K V, Rajendra T K, Lal A K and Lakhotia S C 2000 Omega speckles - a novel class of nuclear speckles containing hnRNPs associated with noncoding hsr-omega RNA in *Drosophila*; *J. Cell Sci.* **113** 3485–3497

Pratt W B 1998 The hsp90-based chaperone system: involvement in signal transduction from a variety of hormone and growth factor receptors; *Proc. Soc. Exp. Biol. Med.* **217** 420–434

Rane M J, Pan Y, Singh S, Powell D W, Wu R, Cummins T, Chen Q, McLeish K R and Klein J B 2003 Heat shock protein 27 controls apoptosis by regulating Akt activation; *J. Biol. Chem.* **278** 27828–27835

Ravagnan L, Gurbuxani S, Susin S A, Maisse C, Daugas E, Zamzami N, Mak T, Jaattela M, Penninger J M, Garrido C and Kroemer G 2001 Heat-shock protein 70 antagonizes apoptosis-inducing factor; *Nat. Cell. Biol.* **3** 839–843

Regnier C H, Song H Y, Gao X, Goeddel D V, Cao Z and Rothe M 1997 Identification and characterization of an IkappaB kinase; *Cell* **90** 373–383

Riedl S J and Shi Y 2004 Molecular mechanisms of caspase regulation during apoptosis; *Nat. Rev. Mol. Cell Biol.* **5** 897–907

Romashkova J A and Makarov S S 1999 NF-kappaB is a target of AKT in anti-apoptotic PDGF signalling; *Nature (London)* **401** 86–90

Rutherford S L and Lindquist S 1998 Hsp90 as a capacitor for morphological evolution; *Nature (London)* **396** 336–342

Saleh A, Srinivasula S M, Balkir L, Robbins P D and Alnemri E S 2000 Negative regulation of the Apaf-1 apoptosome by Hsp70; *Nat. Cell. Biol.* **2** 476–483

Samali A and Orrenius S 1998 Heat shock proteins: regulators of stress response and apoptosis; *Cell Stress Chaperones* **3** 228–236

Samali A, Cai J, Zhivotovsky B, Jones D P and Orrenius S 1999 Presence of a pre-apoptotic complex of pro-caspase-3, Hsp60 and Hsp10 in the mitochondrial fraction of jurkat cells; *EMBO J.* **18** 2040–2048

Sarkar S and Lakhotia S C 2005 The Hsp60C gene in the 25F cytogenetic region in *Drosophila melanogaster* is essential for tracheal development and fertility; *J. Genet.* **84** 265–281

Sarkar S, Arya R and Lakhotia S C 2006 Chaperonins: In life and death; in *Stress Response: A Molecular Biology Approach* (eds) A S Sreedhar and U K Srinivas (Kerala, India: Research Signpost) pp 43–60

Sato S, Fujita N and Tsuruo T 2000 Modulation of Akt kinase activity by binding to Hsp90; *Proc. Natl. Acad. Sci. USA* **97** 10832–10837

Schlesinger M J, Ashburner M and Tissiers A 1982 *Heat shock proteins: from bacteria to man*; (New York: Cold Spring Harbor Laboratory Press)

Schwerk C and Schulze-Osthoff K 2005 Regulation of apoptosis by alternative pre-mRNA splicing; *Mol. Cell* **19** 1–13

Screaton G and Xu X N 2000 T cell life and death signalling via TNF-receptor family members; *Curr. Opin. Immunol.* **12** 316–322

Sengupta S and Lakhotia S C 2006 Altered Expressions of the Noncoding hsromega Gene Enhances poly-Q-Induced Neurotoxicity in *Drosophila*; *RNA Biol.* **3** 28–35

Shan Y X, Liu T J, Su H F, Samsamshariat A, Mestril R and Wang P H 2003 Hsp10 and Hsp60 modulate Bcl-2 family and mitochondria apoptosis signaling induced by doxorubicin in cardiac muscle cells; *J. Mol. Cell. Cardiol.* **35** 1135–1143

Shchors K, Yehiely F, Kular R K, Kotlo K U, Brewer G and Deiss L P 2002 Cell death inhibiting RNA (CDIR) derived from a 3'-untranslated region binds AUF1 and heat shock protein 27; *J. Biol. Chem.* **277** 47061–47072

Simon M M, Reikerstorfer A, Schwarz A, Krone C, Luger T A, Jaattela M and Schwarz T 1995 Heat shock protein 70 overexpression affects the response to ultraviolet light in murine fibroblasts. Evidence for increased cell viability and suppression of cytokine release; *J. Clin. Invest.* **95** 926–933

Sreedhar A S and Csermely P 2004 Heat shock proteins in the regulation of apoptosis: new strategies in tumor therapy: a comprehensive review; *Pharmacol. Ther.* **101** 227–257

Srivastava P 2004 *Studies on the constitutively expressed members of Hsp60 anf Hsp70 gene families in Drosophila melanogaster*, PhD. Dissertation, Banaras Hindu University, Varanasi, India

Staal S P 1987 Molecular cloning of the akt oncogene and its human homologues AKT1 and AKT2: amplification of AKT1 in a primary human gastric adenocarcinoma; *Proc. Natl. Acad. Sci. USA* **84** 5034–5037

Stankiewicz A R, Lachapelle G, Foo C P, Radicioni S M and Mosser D D 2005 Hsp70 inhibits heat-induced apoptosis upstream of mitochondria by preventing Bax translocation; *J. Biol. Chem.* **280** 38729–38739

Stokoe D, Stephens L R, Copeland T, Gaffney P R, Reese C B, Painter G F, Holmes A B, McCormick F and Hawkins P T 1997 Dual role of phosphatidylinositol-3,4,5-trisphosphate in the activation of protein kinase B; *Science* **277** 567–570

Van Gurp M, Festjens N, van Loo G, Saelens X and Vandenabeele P 2003 Mitochondrial intermembrane proteins in cell death; *Biochem. Biophys. Res. Commun.* **304** 487–497

Veereshwarayya V, Kumar P, Rosen K M, Mestril R and Querfurth H W 2006 Differential effects of mitochondrial heat shock protein 60 and related molecular chaperones to prevent intracellular beta-amyloid-induced inhibition of complex IV and limit apoptosis; *J. Biol. Chem.* **281** 29468–29478

Volker U, Mach H, Schmid R and Hecker M 1992 Stress proteins and cross-protection by heat shock and salt stress in Bacillus subtilis; *J. Gen. Microbiol.* **138** 2125–2135

Volloch V, Gabai V L, Rits S and Sherman M Y 1999 ATPase activity of the heat shock protein hsp72 is dispensable for its effects on dephosphorylation of stress kinase JNK and on heat-induced apoptosis; *FEBS Lett.* **461** 73–76

Vydra N, Malusecka E, Jarzab M, Lisowska K, Glowala-Kosinska M, Benedyk K, Widlak P, Krawczyk Z and Widlak W 2006 Spermatocyte-specific expression of constitutively active heat shock factor 1 induces HSP70i-resistant apoptosis in male germ cells; *Cell Death Differ.* **13** 212–222

Wang C Y, Mayo M W, Korneluk R G, Goeddel D V and Baldwin A S, Jr. 1998 NF-kappaB antiapoptosis: induction of TRAF1 and TRAF2 and c-IAP1 and c-IAP2 to suppress caspase-8 activation; *Science* **281** 1680–1683

Warrick J M, Paulson H L, Gray-Board G L, Bui Q T, Fischbeck K H, Pittman R N and Bonini N M 1998 Expanded polyglutamine protein forms nuclear inclusions and causes neural degeneration in *Drosophila*; *Cell* **93** 939–949

White K, Tahaoglu E and Steller H 1996 Cell killing by the *Drosophila* gene reaper; *Science* **271** 805–807

Willis S, Day C L, Hinds M G and Huang D C 2003 The Bcl-2-regulated apoptotic pathway; *J. Cell Sci.* **116** 4053–4056

Wyllie A H, Kerr J F and Currie A R 1980 Cell death: the significance of apoptosis; *Int. Rev. Cytol.* **68** 251–306

Xanthoudakis S, Roy S, Rasper D, Hennessey T, Aubin Y, Cassady R, Tawa P, Ruel R, Rosen A and Nicholson D W 1999 Hsp60 accelerates the maturation of pro-caspase-3 by upstream activator proteases during apoptosis; *EMBO J.* **18** 2049–2056

Xia W, Voellmy R and Spector N L 2000 Sensitization of tumor cells to fas killing through overexpression of heat-shock transcription factor 1; *J Cell Physiol* **183** 425–431

Yan N and Shi Y 2005 Mechanisms of apoptosis through structural biology; *Annu. Rev. Cell Dev. Biol.* **21** 35–56

Yang X, Khosravi-Far R, Chang H Y and Baltimore D 1997 Daxx, a novel Fas-binding protein that activates JNK and apoptosis; *Cell* **89** 1067–1076

Zha J, Harada H, Yang E, Jockel J and Korsmeyer S J 1996 Serine phosphorylation of death agonist BAD in response to survival factor results in binding to 14-3-3 not BCL-X(L); *Cell* **87** 619–628

Zhang R, Luo D, Miao R, Bai L, Ge Q, Sessa W C and Min W 2005 Hsp90-Akt phosphorylates ASK1 and inhibits ASK1-mediated apoptosis; *Oncogene* **24** 3954–3963

Zhang S Q, Kovalenko A, Cantarella G and Wallach D 2000 Recruitment of the IKK signalosome to the p55 TNF receptor: RIP and A20 bind to NEMO (IKKgamma) upon receptor stimulation; *Immunity* **12** 301–311

Zou H, Li Y, Liu X and Wang X 1999 An APAF-1.cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9; *J. Biol. Chem.* **274** 11549–11556

*e*Publication: 22 March 2007

EXHIBIT 3

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 2066–2070, March 1994
Biochemistry

A gene encoding a DnaK/hsp70 homolog in *Escherichia coli*

(chaperone/heat shock cognate protein/ferredoxin operon)

BRENT L. SEATON AND LARRY E. VICKERY*

Department of Physiology and Biophysics, University of California, Irvine, CA 92717

Communicated by Helmut Beinert, November 29, 1993 (received for review May 6, 1993)

ABSTRACT **Eukaryotic organisms have been shown to have multiple forms of hsp70-class stress-related proteins, but only a single family member, DnaK, has been found in prokaryotes. We report here the identification of a heat shock cognate gene, designated *hsc*, in *Escherichia coli*. The amino acid sequence deduced from *hsc* predicts a 65,647-Da polypeptide having 41% sequence identity with DnaK from *E. coli*, and overexpression produces a protein (Hsc66) with properties similar to DnaK. In contrast to *dnaK*, however, the *hsc* gene lacks a consensus heat shock promoter sequence, and expression is not induced by elevated temperature. The *hsc* gene is located near 54 min on the physical map, immediately upstream of the *fdx* gene, which encodes a [2Fe–2S] ferredoxin; evidence is presented that the *hsc* and *fdx* genes make up a bicistronic operon in which expression of the ferredoxin is coupled to that of Hsc66. The function of Hsc66 is not known, but the coregulation of its expression with that of ferredoxin suggests the possibility of a specific role in association with the ferredoxin protein.**

The 70-kDa heat shock proteins (hsp70) and their cognates (hsc70) make up a ubiquitous, multigene family of highly conserved proteins, which are involved in diverse protein–protein interactions (reviewed in ref. 1). They are important under normal conditions as well as during stress and have been implicated in a variety of processes including stabilization of protein-folding intermediates (2), protein assembly and disassembly (3), protein secretion (4, 5), and protein degradation (6). Eukaryotic organisms have been found to contain multiple hsp70 family members; for example, nine distinct proteins are produced in *Saccharomyces cerevisiae* (3, 7), six have been identified in *Drosophila* (8), and at least eight have been described in mammals (9). In contrast, only a single hsp70-class protein, DnaK, has been reported in prokaryotes. The most extensively characterized of these hsp70 proteins is the DnaK protein from *Escherichia coli*. DnaK plays a role in the heat shock response (10, 11), but it is expressed at levels of ≈1% of the cell protein and performs important cellular functions under nonstress conditions (12–14). Hybridization analyses in *E. coli* have revealed only one gene, *dnaK*, located near 0.3 min on the genetic linkage map (15, 16). The presence of a single gene encoding a "stress 70-type" protein in prokaryotes would seem to suggest that all members of the multigene eukaryote hsp70 family evolved from a single DnaK-like ancestral protein (cf. refs. 17 and 18).

We report here the identification of a second hsp70-related gene in *E. coli*. The gene,† designated *hsc*, encodes a protein of ≈66 kDa (Hsc66), which shows ≈40% sequence identity with DnaK and other hsp70-class proteins. The *hsc* gene is found near 54 min on the *E. coli* chromosome and is located immediately upstream of the *fdx* gene, which encodes a [2Fe–2S] ferredoxin (19, 20). These genes appear to make up a bicistronic operon in which expression of Hsc66 and ferredoxin is coregulated.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

EXPERIMENTAL PROCEDURES

General Methods. Expression of Hsc66 was carried out in *E. coli* strain MZ-1 (21). Sequencing, bacterial transformation, and oligonucleotide purification were carried out as described by Sambrook *et al.* (22), and β-galactosidase activities were determined as described by Miller (23). SDS/PAGE was carried out according to Laemmli (24). Western immunoblotting was carried out by the method of Towbin *et al.* (25) using enhanced chemiluminescence detection (Amersham).

Plasmids. The plasmid p66-Fdx, used to overexpress Hsc66, contained the *hsc* and *fdx* genes and flanking regions derived from clone DT10 originally isolated from an *E. coli* B genomic library (19, 20).‡ An *Eco*RI fragment containing the insert was cloned into pBS(+/−) (Stratagene) to yield the plasmid pDT10 and into pΔblue (27) to yield the plasmid pΔDT10. Five hundred and four base pairs of 5′ flanking DNA were deleted from pΔDT10 by digesting with *Nco* I and *Hin*dIII, and the overhangs were filled in with Klenow DNA polymerase and ligated. The resulting plasmid, p66-Fdx, contained the *hsc* gene, including 188 bp of 5′ flanking DNA, under control of the λ p_L promoter.

Plasmid p66-Lac, used for analyses of Hsc66 expression, was constructed by amplifying the region of pDT10 containing the N-terminal nine amino acids of Hsc66 and 690 bp of upstream sequence. PCR primers were constructed such that the 5′ end of the PCR product would contain an *Eco*RI site and the 3′ end would contain an *Sma* I site to allow fusion of the *hsc* coding sequence in-frame with a *lacZ* coding sequence in pMLB1034 (28). The PCR product was digested with *Eco*RI and *Sma* I, and the 730-bp fragment was ligated to pMLB1034 that had been digested with *Eco*RI and *Sma* I.

Plasmid pFdx-Lac, used for promoter analyses, was constructed using PCR to amplify a region from pDT10 including ≈900 bp upstream of the *fdx* coding region and the bases encoding the first 11 amino acids of ferredoxin. PCR primers were designed such that the 3′ end of the amplified sequence would contain a *Bam*HI site. The PCR product was cleaved with *Acc* I, at a site located 91 bp upstream of the *fdx* coding sequence, and *Bam*HI, and the resulting 124-bp fragment was ligated to the 2.4-kb *Eco*RI–*Acc* I fragment of pDT10 containing the remaining upstream sequences of DT10. This fragment was then ligated to the 3.1-kb *Eco*RI–*Bam*HI fragment of pMLB1034. The resulting pFdx-Lac contains 2.5 kb of sequence upstream of the *fdx* gene followed by the sequence encoding the first 11 amino acids of ferredoxin fused in-frame to β-galactosidase beginning at codon 8. Upstream deletion derivatives of pFdx-Lac shown in Fig. 5 were prepared by

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. U05338).

‡The *hsc* and *fdx* genes are also present in *E. coli* K12 and are located in λ clones 7F8 and 5E10 (20) in the miniset library isolated by Kohara *et al.* (26).

digesting pFdx-Lac with the appropriate restriction enzymes, isolating the large vector fragment by agarose electrophoresis, and ligating the purified deletion construct.

Protein Expression and Purification. MZ-1 cells transformed with plasmid p66-Fdx were grown in Terrific broth (22) to $OD_{600} \approx 0.5$, induced by heating to 42°C for 2 hr, and subsequently grown overnight at 37°C. Cells were harvested by centrifugation and disrupted by French press. Protein extracts were fractionated by anion-exchange chromatography and molecular sieving chromatography on Sephacryl S-300. The description of a more complete purification procedure will be published (L. W. Goodman and L.E.V.).

Primer Extension. Primer extension analysis was carried out using a synthetic oligonucleotide complementary to nucleotides 22–41 of the *hsc* coding sequence. The oligonucleotide was labeled at the 5′ end using [γ-^{32}P]ATP and hybridized to 50 μg of total RNA isolated from *E. coli* strain DH5α previously transformed with pDT10. RNA was isolated from late-logarithmic phase cells by breaking with glass beads in the presence of hot (65°C), water-saturated phenol. The mixture was vortexed for 30 s followed by incubation at 65°C for 30 s; this cycle was repeated twice. The aqueous phase was extracted with chloroform and precipitated with ethanol. RNA samples were resuspended in RNase-free water and treated with RNase-free DNase. Hybridizations and primer extensions were performed using the Promega primer extension kit following the included protocol. After extension with reverse transcriptase, samples were digested with RNase A for 30 min at 37°C and extracted with phenol. RNA·DNA hybrids were precipitated with ethanol, denatured, and subjected to electrophoresis on a 6% polyacrylamide sequencing gel. Unlabeled oligonucleotide served as a primer for the sequencing ladder comparison.

RESULTS AND DISCUSSION

Identification of the *hsc* Gene. An open reading frame of 1848 bp encoding a possible hsp70-class protein was detected during sequencing of DNA in the 5′ flanking region of the *fdx* gene of *E. coli* (Fig. 1). Analysis of codon usage with the *E. coli* codon bias (29) shows a clear preference for this reading frame, and a sequence resembling a Shine–Dalgarno sequence for ribosome binding and translation (30) is found immediately upstream of the initiation AUG. Sequence showing similarity to a −10 consensus promoter sequence is also observed from −63 to −68 bp, but no −35 region consensus sequence (31) is apparent.

The translated DNA sequence of this reading frame predicts a polypeptide of 616 amino acids with a molecular mass of 65,647 Da. A search of the GenBank data base using the deduced amino acid sequence indicated that the predicted protein showed homology to prokaryotic and eukaryotic hsp70-class stress proteins, and among the 100 proteins exhibiting the highest similarity scores, all were either heat shock or heat shock cognate proteins. Because of the apparent lack of a heat shock promoter consensus sequence in the 5′ flanking region of the gene (cf. ref. 7; see also below), we considered the gene product to be a heat shock cognate protein and designated the gene *hsc* and the predicted protein Hsc66.

Comparison of Hsc66 with DnaK. The protein exhibiting the highest degree of sequence similarity to Hsc66 is DnaK of *E. coli*, and a comparison of the amino acid sequences predicted for the two proteins is presented in Fig. 2. The alignment shown yields a sequence identity of 41% and a similarity of 60% over the region of residues 17–616 of Hsc66. Similarities to other hsp70 proteins are also notable: Hsc66 has 36% identity with bovine hsc70 (32) and 39% identity with yeast Ssc1p over the same region (33). The similarities observed are especially notable in the N-terminal two-thirds of the proteins; this is the most highly conserved region in hsp70 proteins and has been identified as an ATPase domain in other forms of hsp70 (34). A number of the conserved residues have been shown to be involved in ATP binding in bovine Hsc70 (35, 36), and Hsc66 residues 208–230 show homology to the ATP binding sites of protein kinases (37). These similarities suggest that, like other hsp70 proteins, Hsc66 may possess ATPase activity. In addition, Thr-212 of Hsc66 aligns with Thr-199 of DnaK, a site of autophosphorylation (38), raising the possibility that Hsc66 may also be subject to regulation by phosphorylation at this position.

Significant differences between the sequences predicted for Hsc66 and DnaK, however, are apparent. It was necessary to introduce several gaps in the sequence of Hsc66 to optimize the alignment. Alignment of the sequences of Hsc66 and DnaK with the structure of the ATPase fragment of bovine Hsc70 (35, 36) suggests that the regions in which gaps were introduced may correspond to residues near the surface of the folded protein; thus, these differences may reflect different surface structural features of Hsc66 compared to DnaK and bovine Hsc70 in those regions. In addition, Hsc66 is predicted to have an N-terminal extension not present in DnaK and lacks 17 C-terminal residues found in DnaK. The 16-residue N-terminal extension of Hsc66, which is not present in *E. coli* DnaK, is unusual because a similar extension is absent from the predicted sequences of DnaK proteins found in other prokaryotes. Some eukaryotic forms of hsp70 contain N-terminal extensions, which function in targeting to, or retention in, the endoplasmic reticulum or mitochondria (for review, see ref. 38), but these do not show sequence similarity to the N terminus of Hsc66. Moreover, the N-terminal sequence of Hsc66 does not show homology to known signal sequences of membrane-bound or periplasmic proteins of *E. coli*, and its role remains to be determined. The divergence between the amino acid sequences of Hsc66 and DnaK in the C-terminal region is similar to the variability observed in other hsp70 proteins. The C-terminal domain is believed to be involved in protein recognition (37–39), and the differences observed suggest that Hsc66 is likely to interact with different target protein(s) within the cell.

Expression of Hsc66. To establish the identity of the *hsc* gene product, the plasmid p66-Fdx, containing 188 bp of 5′ flanking DNA and the putative coding region, was constructed to overexpress Hsc66 in MZ-1 cells. Fractionation of extracts from induced cells revealed a major band of ≈66 kDa, which was partially purified by anion-exchange and gel-filtration chromatography. Fig. 3 *Left* shows the preparation following SDS/PAGE and blotting to a poly(vinylidene difluoride) membrane. The major band of ≈66 kDa was subjected to N-terminal amino acid sequencing for nine cycles and was identified as Hsc66; the first residue of the mature protein was found to be alanine, indicating that the formylmethionine was removed as found for other proteins in *E. coli* which have alanine as the penultimate N-terminal residue (40). DnaK, which has properties similar to Hsc66, copurified in the preparation and is visible as a minor band that migrates with an apparent molecular mass of ≈75 kDa, an anomaly previously reported (41).

Western immunoblot analyses were carried out on the partially purified preparation to test the relatedness of Hsc66 and DnaK. Antisera to DnaK from three rabbits were separately tested, and in no case was cross-reactivity with Hsc66 observed. As shown in Fig. 3 *Right*, DnaK is readily detected in the preparation by each antiserum, whereas Hsc66, although present in larger amounts, is not detected, suggesting that the major epitopes present in DnaK are not conserved in Hsc66.

Characterization of the *hsc–fdx* Operon. The 5′ flanking region of the *hsc* gene does not contain sequences resembling the consensus sequences found in heat shock promoters (-CnCccTTGAA- in the −35 region and -CCCCATnT- in the −10 region), which are recognized by the heat shock σ^{32} factor (7).

```
-180                -160                -140                -120                -100
TCAAACGTGT GAAAAAGATG TTTGATACCC GCCATCAGTT GATGGTTGAA CAGTTAGACA ACGAGACGTG GGACGCGGCG GCGGATACCG
        -80          (-10)  -60 ➥              -40                 -20          S-D
TGCGTAAGCT GCGTTTTCTC GATAAACTGC GAAGCAGTGC CGAACAACTC GAAGAAAAAC TGCTCGATTT TTAATTTCTG GAAGCTAAAC
1                           30                                  60
ATG GCC TTA TTA CAA ATT AGT GAA CCT GGT TTG AGT GCC GCG CCG CAT CAG CGT CGT CTG GCG GCC GGT ATT GAC
Met Ala Leu Leu Gln Ile Ser Glu Pro Gly Leu Ser Ala Ala Pro His Gln Arg Arg Leu Ala Ala Gly Ile Asp
                    90                                  120                                 150
CTG GGC ACA ACC AAC TCG CTG GTG GCG ACA GTG CGC AGC GGT CAG GCC GAA ACG TTA GCC GAT CAT GAA GGC CGT
Leu Gly Thr Thr Asn Ser Leu Val Ala Thr Val Arg Ser Gly Gln Ala Glu Thr Leu Ala Asp His Glu Gly Arg
                                180                                 210
CAC CTG CTG CCA TCT GTT GTT CAC TAT CAA CAG CAA GGG CAT TCG GTG GGT TAT GAC GCG CGT ACT AAT GCA GCG
His Leu Leu Pro Ser Val Val His Tyr Gln Gln Gln Gly His Ser Val Gly Tyr Asp Ala Arg Thr Asn Ala Ala
                    240                                 270                                 300
CTC GAT ACC GCC AAC ACA ATT AGT TCT GTT AAA CGC CTG ATG GGA CGC TCG CTG GCT GAT ATC CAG CAA CGC TAT
Leu Asp Thr Ala Asn Thr Ile Ser Ser Val Lys Arg Leu Met Gly Arg Ser Leu Ala Asp Ile Gln Gln Arg Tyr
                                330                                 360
CCG CAT CTG CCT TAT CAA TTC CAG GCC AGC GAA AAC GGC CTG CCG ATG ATT GAA ACG GCG GCG GGG CTG CTG AAC
Pro His Leu Pro Tyr Gln Phe Gln Ala Ser Glu Asn Gly Leu Pro Met Ile Glu Thr Ala Ala Gly Leu Leu Asn
                    390                                 420                                 450
CCG GTG CGC GTT TCT GCG GAC ATC CTC AAA GCA CTG GCG GCG CGG GCA ACT GAA GCC CTG GCA GGC GAG CTG GAT
Pro Val Arg Val Ser Ala Asp Ile Leu Lys Ala Leu Ala Ala Arg Ala Thr Glu Ala Leu Ala Gly Glu Leu Asp
                                480                                 510
GGT GTA GTT ATC ACC GTT CCG GCG TAC TTT GAC GAT GCC CAG CGT CAG GGC ACC AAA GAC GCG GCG CGT CTG GCG
Gly Val Val Ile Thr Val Pro Ala Tyr Phe Asp Asp Ala Gln Arg Gln Gly Thr Lys Asp Ala Ala Arg Leu Ala
                    540                                 570                                 600
GGC CTT CAC GTC CTG CGC TTA CTT AAC GAA CCG ACC GCT GCG GCT ATC GCC TAC GGG CTG GAT TCC GGT CAG GAA
Gly Leu His Val Leu Arg Leu Leu Asn Glu Pro Thr Ala Ala Ala Ile Ala Tyr Gly Leu Asp Ser Gly Gln Glu
                                630                                 660
GGC GTG ATC GCC GTT TAT GAC CTC GGT GGC GGG ACG TTT GAT ATT TCC ATT CTG CGC TTA AGT CGC GGC GTG TTT
Gly Val Ile Ala Val Tyr Asp Leu Gly Gly Gly Thr Phe Asp Ile Ser Ile Leu Arg Leu Ser Arg Gly Val Phe
                    690                                 720                                 750
GAA GTG CTG GCA ACC GGC GGT GAT TCC GCG CTC GGC GGC GAT GAT TTC GAC CAT CTG CTG GCG GAT TAC ATT CGC
Glu Val Leu Ala Thr Gly Gly Asp Ser Ala Leu Gly Gly Asp Asp Phe Asp His Leu Leu Ala Asp Tyr Ile Arg
                                780                                 810
GAG CAG GCG GGC ATT CCT GAT CGT AGC GAT AAC CGC GTT CAG CGT GAA CTG CTG GAT GCC GCC ATT GCA GCC AAA
Glu Gln Ala Gly Ile Pro Asp Arg Ser Asp Asn Arg Val Gln Arg Glu Leu Leu Asp Ala Ala Ile Ala Ala Lys
                    840                                 870                                 900
ATC GCG CTG AGC GAT GCG GAC TCC GTG ACC GTT AAC GTT GCG GGC TGG CAG GGC GAA ATC AGC CGT GAA CAA TTC
Ile Ala Leu Ser Asp Ala Asp Ser Val Thr Val Asn Val Ala Gly Trp Gln Gly Glu Ile Ser Arg Glu Gln Phe
                                930                                 960
AAT GAA CTG ATC GCG CCA CTG GTA AAA CGA ACC TTA CTG GCT TGT CGT CGC GCG CTG AAA GAC GCG GGT GTA GAA
Asn Glu Leu Ile Ala Pro Leu Val Lys Arg Thr Leu Leu Ala Cys Arg Arg Ala Leu Lys Asp Ala Gly Val Glu
                    990                                1020                                1050
GCT GAT GAA GTG CTG GAA GTG GTG ATG GTG GGC GGT TCT ACT CGC GTG CCG CTG GTG CGT GAA CGG GTA GGC GAA
Ala Asp Glu Val Leu Glu Val Val Met Val Gly Gly Ser Thr Arg Val Pro Leu Val Arg Glu Arg Val Gly Glu
                               1080                                1110
TTT TTC GGT CGT CCA CCG CTG ACT TCC ATC GAC CCG GAT AAA GTC GTC GCT ATT GGC GCG GCG ATT CAG GCG GAT
Phe Phe Gly Arg Pro Pro Leu Thr Ser Ile Asp Pro Asp Lys Val Val Ala Ile Gly Ala Ala Ile Gln Ala Asp
                   1140                                1170                                1200
ATT CTG GTG GGT AAC AAG CCA GAC AGC GAA ATG CTG CTG CTT GAT GTG ATC CCA CTG TCG CTG GGC CTC GAA ACG
Ile Leu Val Gly Asn Lys Pro Asp Ser Glu Met Leu Leu Leu Asp Val Ile Pro Leu Ser Leu Gly Leu Glu Thr
                               1230                                1260
ATG GGC GGT CTG GTG GAG AAA GTG ATT CCG CGT AAT ACC ACT ATT CCG GTG GCC CGC GCT CAG GAT TTC ACC ACC
Met Gly Gly Leu Val Glu Lys Val Ile Pro Arg Asn Thr Thr Ile Pro Val Ala Arg Ala Gln Asp Phe Thr Thr
                   1290                                1320                                1350
TTT AAA GAT GGT CAG ACG GCG ATG TCT ATC CAT GTA ATG CAG GGT GAG CGC GAA CTG GTG CAG GAC TGC CGC TCA
Phe Lys Asp Gly Gln Thr Ala Met Ser Ile His Val Met Gln Gly Glu Arg Glu Leu Val Gln Asp Cys Arg Ser
                               1380                                1410
CTG GCG CGT TTT GCG CTG CGT GGT ATT CCG GCG CTA CCG GCT GGC GGT GCG CAT ATT CGC GTG ACG TTC CAG GTC
Leu Ala Arg Phe Ala Leu Arg Gly Ile Pro Ala Leu Pro Ala Gly Gly Ala His Ile Arg Val Thr Phe Gln Val
                   1440                                1470                                1500
GAT GCC GAC GGT CTT TTG AGC GTG ACG GCG ATG GAG AAA TCC ACC GGC GTT GAG GCG TCT ATT CAG GTC AAA CCG
Asp Ala Asp Gly Leu Leu Ser Val Thr Ala Met Glu Lys Ser Thr Gly Val Glu Ala Ser Ile Gln Val Lys Pro
                               1530                                1560
TCT TAC GGT CTG ACT GAC AGC GAA ATC GCT TCG ATG ATC AAA GAC TCA ATG AGC TAT GCC GAG CAG GAC GTA AAA
Ser Tyr Gly Leu Thr Asp Ser Glu Ile Ala Ser Met Ile Lys Asp Ser Met Ser Tyr Ala Glu Gln Asp Val Lys
                   1590                                1620                                1650
GCC CGA ATG CTG GCA GAA CAA AAA GTA GAA GCG GCG CGT GTG CTG GAA AGT CTG CAC GGC GCG CTG GCT GCT GAT
Ala Arg Met Leu Ala Glu Gln Lys Val Glu Ala Ala Arg Val Leu Glu Ser Leu His Gly Ala Leu Ala Ala Asp
                               1680                                1710
GCC GCG CTG TTA AGC GCC GCA GAA CGT CAG GTC ATT GAC GAT GCT GCC GCT CAC CTG AGT GAA GTG GCG CAG GGC
Ala Ala Leu Leu Ser Ala Ala Glu Arg Gln Val Ile Asp Asp Ala Ala Ala His Leu Ser Glu Val Ala Gln Gly
                   1740                                1770                                1800
GAT GAT GTT GAC GCC ATC GAA AAA GCG ATT AAA AAC GTA GAC AAA CAA ACC CAG GAT TTC GCC GCT CGC CGC ATG
Asp Asp Val Asp Ala Ile Glu Lys Ala Ile Lys Asn Val Asp Lys Gln Thr Gln Asp Phe Ala Ala Arg Arg Met
                               1830                1848        Fdx →
GAC CAG TCG GTT CGT CGT GCG CTG AAA GGC CAT TCC GTG GAC GAG GTT TAA T ATG CCA AAG ATT GTT ATT TTG
Asp Gln Ser Val Arg Arg Ala Leu Lys Gly His Ser Val Asp Glu Val  *    Met Pro Lys Ile Val Ile Leu
```

FIG. 1. Nucleotide sequence and deduced amino acid sequence of the *hsc* gene. DNA sequence numbering begins with the predicted initiator methionine of Hsc66; the initiator methionine for the *fdx* gene is indicated by Fdx. Possible regulatory sequences are underlined (−10, promoter sequence; S-D, Shine–Dalgarno sequence), and the proposed site of transcriptional initiation (see Fig. 3) is indicated by an arrow (→).

To test whether other sequences present might function in heat shock induction of the gene, we used the vector p66-Lac, which contains the *lacZ* gene fused in-frame with bases encoding the first nine amino acids of Hsc66 together with 690 bp of 5′ flanking DNA; control of expression of the chimeric gene is thus under control of *hsc* promoter sequences. This plasmid was introduced into *E. coli* JM109 cells, and β-galactosidase activity of cell extracts was determined before and after subjecting cultures to heat shock. No increase in β-galactosidase activity was observed following a shift to 46°C or 51°C for up to 30 min (data not shown), suggesting that Hsc66 is not induced by heat shock and is subject to other control mechanisms.

The observation that only a single base separates the termination codon for *hsc* and the initiation codon for *fdx* suggests that the two genes might function as a bicistronic operon. To determine whether the genes are cotranscribed, the plasmid

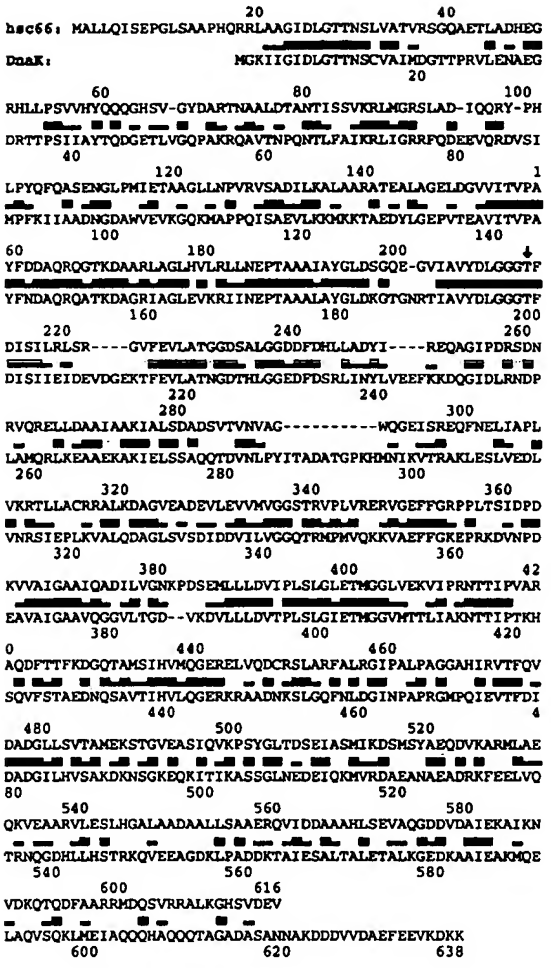


FIG. 2. Comparison of the deduced amino acid sequences of *E. coli* Hsc66 and DnaK. Amino acid identities are denoted by a thick line, and similarities are denoted by a thinner line.

pFdx-Lac was constructed (Fig. 4). This plasmid encodes a ferredoxin–β-galactosidase fusion protein under the control of promoter elements immediately upstream of the *fdx* gene as well as those in the 5′ flanking region of the *hsc* gene. Deletion derivatives of pFdx-Lac were made using unique restriction sites at varying distances from the 5′ end of the insert. Plasmids were introduced into *E. coli* JM109 cells, and the cells were grown to midlogarithmic phase for determination of β-galactosidase activity levels. Deletion of the *Eco*RI–*Hin*dIII region of the upstream sequence reduced β-galactosidase activity by only 7%, whereas deletion of the *Eco*RI–*Nru* I region reduced activity by 94%; no further reduction was observed upon deletion of bases to within 90 bp of the *fdx* initiation codon. These findings suggest that under the growth conditions used expression of the *fdx* gene is primarily regulated by promoter sequences between the *Hin*dIII and *Nru* I sites.

A derivative of plasmid pFdx-Lac was also constructed in which an 8-bp linker was inserted into the *Nru* I site found at position 750 of the *hsc* coding sequence. This insertion causes a shift of the reading frame of *hsc* and introduces a termination codon 28 codons after the site of linker insertion. As shown in Fig. 4, this frameshift reduced the β-galactosidase activity of the *fdx–lacZ* fusion ≈7-fold. This finding suggests the possibility that *hsc* and *fdx* are translationally coupled, with translation of the ferredoxin mRNA dependent on translation of the Hsc66 mRNA.

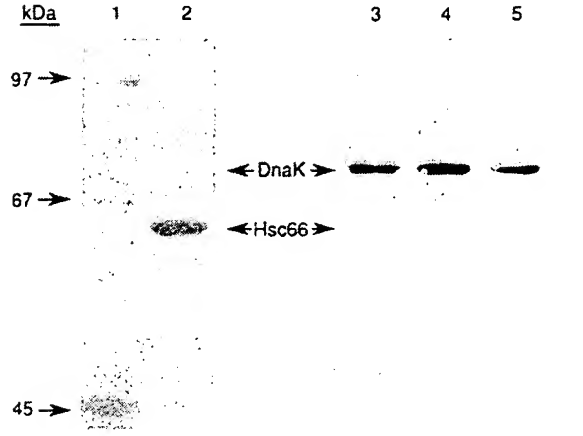


FIG. 3. SDS/PAGE and immunoblot analysis of Hsc66. A partially purified preparation of Hsc66 was subjected to SDS/PAGE in a 10% gel, and the gel was blotted to a poly(vinylidene difluoride) membrane and stained with Coomassie blue to visualize proteins. Lane 1, molecular mass markers; lanes 2–5, ≈3 μg of total protein. The membrane containing lanes 1 and 2 shows Coomassie blue-stained protein bands. The membranes containing lanes 3–5 were individually probed with antisera to *E. coli* DnaK obtained from three rabbits (114, 115B, and 116C, respectively) provided by Graham Walker (Massachusetts Institute of Technology); cross-reacting proteins were detected using a peroxidase-conjugated goat anti-rabbit second antibody and luminol chemiluminescence exposure to autoradiography film.

To determine the 5′ end of the *hsc–fdx* transcript(s), primer extension reactions were performed using total RNA isolated from *E. coli* cells transformed with plasmid pDT10. Reactions were primed with an oligonucleotide complementary to nucleotides 22–41 within the coding sequence of *hsc* or to nucleotides 43–60 within the coding sequence of *fdx*. The results using the primer within the *hsc* gene showed a single major transcript starting 57 bp upstream of the initiation ATG (Fig. 5). The DNA sequence upstream at positions −68 to −63 (TAAACT) shows similarity and spacing to that of −10 promoter sequences; no sequence showing similarity to the −35 consensus promoter sequence (TTGACG; ref. 31), however, is apparent. Using the primer within the *fdx* gene no single, major transcriptional start site was observed within the resolution of the sequencing gel. Instead, a band of moderate intensity (≈1/10th as intense as the *hsc* transcript) was seen, which corresponded to a transcriptional start site 21 bp upstream of the *fdx* initiation ATG; multiple sites of weak intensity were also observed along the full

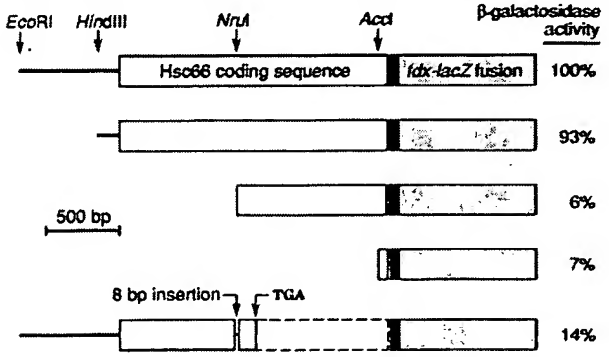


FIG. 4. Promoter analysis of the *hsc–fdx* operon. Deletions and insertions within the plasmid pFdx-Lac were made at the indicated restriction sites. β-Galactosidase activities were measured as described in *Experimental Procedures* and are reported as percent activity of the parent plasmid (100% = 4227 Miller units/mg of protein).

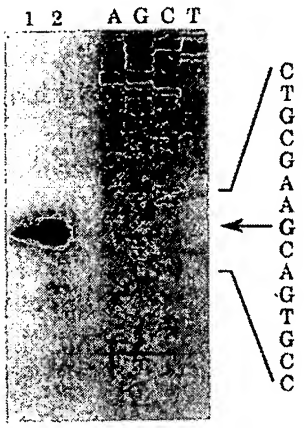


FIG. 5. Primer extension analysis of the transcriptional start site of the *hsc* gene. A [γ-^{32}P]ATP-labeled primer complementary to the 5′ end of the *hsc* gene was hybridized to 50 μg of total RNA isolated from *E. coli* strain DH5α transformed with pDT10 and was extended with reverse transcriptase. Products were analyzed by electrophoresis on a 6% polyacrylamide sequencing gel. Samples in lanes 1 and 2 represent reactions in which the hybridization products were digested with RNase A (lane 1) or were left undigested (lane 2). Lanes A, G, C, and T are products of sequencing reactions using the same oligonucleotide as primer. The sequence of the antisense strand is shown on the right, and the putative start site, G at position −57, is indicated by the arrow.

length of the resolving gel (data not shown). These results, in accordance with the deletion studies described above, suggest that under the conditions used the primary transcript is initiated upstream of the *hsc* gene.

The finding that the *hsc* and *fdx* genes are coregulated suggests that they make up a bicistronic operon. While no complete consensus promoter sequence can be identified within the 5′ flanking region of the *hsc* gene, expression levels observed with the Hsc66– and ferredoxin–β-galactosidase fusions and the primer extension results are consistent with a primary site of transcriptional initiation 57 bp upstream of the Hsc66 coding sequence. In addition, the presence of a sequence resembling a ρ-independent termination signal 120 bp downstream of the *fdx* gene and the absence of open reading frames in this region (19) suggest that no additional genes are encoded in the operon. Additional studies are needed, however, to define the transcript and its regulation.

The Function of Hsc66. Hsp70 proteins participate in a variety of processes involving protein folding, protein assembly and disassembly, protein secretion, and protein degradation. The function of Hsc66 in *E. coli* is not known, but the apparent lack of induction in response to heat shock suggests a role or role(s) in normal cell metabolism as opposed to stress conditions. The finding that the expression of ferredoxin is coregulated with that of Hsc66 suggests the possibility that Hsc66 may function in some way with the ferredoxin protein, perhaps assisting in protein folding or assembly of the iron–sulfur cluster.

Note Added in Proof. The *hsc* gene was independently discovered in *E. coli* K12 by Kawula and Lelivelt (42).

We are grateful to G. Wesley Hatfield and John Keener for helpful discussions, and we thank Graham Walker for providing antisera to *E. coli* DnaK. This work was supported by grants from the National Institutes of Health.

1. McKay, D. B. (1993) *Adv. Prot. Chem.* **44**, 67–98.
2. Langer, T., Lu, C., Echols, H., Flanagan, J., Hayer, M. K. & Hartl, F. U. (1992) *Nature (London)* **356**, 683–689.
3. Lindquist, S. & Craig, E. A. (1988) *Annu. Rev. Genet.* **22**, 631–677.
4. Chirico, W. J., Waters, M. G. & Blobel, G. (1988) *Nature (London)* **332**, 805–810.
5. Deshaies, R. J., Koch, B. D., Werner-Washburene, M., Craig, E. A. & Schekman, R. (1988) *Nature (London)* **332**, 800–805.
6. Chiang, H.-L., Terlecky, S. R., Plant, C. P. & Dice, J. F. (1989) *Science* **246**, 382–385.
7. Nover, L. (1991) *Heat Shock Response* (CRC, Boca Raton, FL).
8. Craig, E. A., Ingolia, T. D. & Manseau, L. J. (1983) *Dev. Biol.* **99**, 418–426.
9. Arrigo, A.-P. & Welch, W. J. (1987) *J. Biol. Chem.* **262**, 15359–15369.
10. Straus, D. B., Walter, W. A. & Gross, C. A. (1989) *Genes Dev.* **3**, 2003–2010.
11. Straus, D., Walter, W. & Gross, C. (1990) *Genes Dev.* **4**, 2202–2209.
12. Bukau, B. & Walker, G. C. (1989) *J. Bacteriol.* **171**, 2337–2346.
13. Bukau, B. & Walker, G. C. (1989) *J. Bacteriol.* **171**, 6030–6038.
14. Herendeen, S. L., VanBogelen, R. A. & Neidhardt, F. C. (1979) *J. Bacteriol.* **139**, 185–194.
15. Bardwell, J. C. A. & Craig, E. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 848–852.
16. Neidhardt, F. C., VanBogelen, R. A. & Vaughn, V. (1984) *Annu. Rev. Genet.* **18**, 295–329.
17. Gupta, R. S. & Singh, B. (1992) *J. Bacteriol.* **174**, 4594–4605.
18. Hughes, A. L. (1993) *Mol. Biol. Evol.* **10**, 243–255.
19. Ta, D. T. & Vickery, L. E. (1992) *J. Biol. Chem.* **267**, 11120–11125.
20. Ta, D. T., Seaton, B. L. & Vickery, L. E. (1992) *J. Bacteriol.* **174**, 5760–5761.
21. Nagai, K. & Thogersen, H. C. (1984) *Nature (London)* **309**, 810–812.
22. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
23. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab. Press, Plainview, NY).
24. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
25. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.
26. Kohara, Y., Akiyama, K. & Isono, K. (1987) *Cell* **50**, 495–508.
27. Brandt, M. E. & Vickery, L. E. (1992) *Arch. Biochem. Biophys.* **294**, 735–740.
28. Silhavy, T. J., Berman, M. L. & Enquist, L. W. (1984) *Experiments with Gene Fusions* (Cold Spring Harbor Lab. Press, Plainview, NY).
29. Gribskov, M., Devereaux, J. & Burgess, R. R. (1984) *Nucleic Acids Res.* **12**, 539–549.
30. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
31. Hawley, D. K. & McClure, W. R. (1983) *Nucleic Acids Res.* **11**, 2237–2255.
32. DeLuca-Flaherty, C. & McKay, D. B. (1990) *Nucleic Acids Res.* **18**, 5569.
33. Craig, E. A., Kramer, J., Shilling, J., Werner-Washburne, M., Holmes, S., Kosic-Smithers, J. & Nicolet, C. M. (1989) *Mol. Cell. Biol.* **9**, 3000–3008.
34. Chappell, T. G., Konforti, B. B., Schmid, S. L. & Rothman, J. E. (1987) *J. Biol. Chem.* **262**, 746–751.
35. Flaherty, K. M., DeLuca-Flaherty, C. & McKay, D. B. (1990) *Nature (London)* **346**, 623–628.
36. Flaherty, K. M., McKay, D. B., Kabsch, W. & Holmes, K. C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 5041–5045.
37. Hannink, M. & Donoghue, D. J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7894–7898.
38. McCarty, J. S. & Walker, G. C. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 9513–9517.
39. Gething, M.-J. & Sambrook, J. (1992) *Nature (London)* **355**, 33–45.
40. Flinta, C., Persson, B., Jornvall, H. & von Heijne, G. (1986) *Eur. J. Biochem.* **154**, 193–196.
41. Zylicz, M. & Georgopoulos, C. (1984) *J. Biol. Chem.* **259**, 8820–8825.
42. Kawula, T. H. & Lelivelt, M. J. (1994) *J. Bacteriol.* **176**, 610–619.

EXHIBIT 4

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1993 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 268, No. 30, Issue of October 25, pp. 22609-22617, 1993
Printed in U.S.A.

Isolation and Characterization of ClpX, a New ATP-dependent Specificity Component of the Clp Protease of *Escherichia coli**

(Received for publication, April 19, 1993, and in revised form, June 18, 1993)

Diana Wojtkowiak‡, Costa Georgopoulos§, and Maciej Zylicz‡¶

From the ‡Department of Molecular Biology, Division of Biophysics, University of Gdansk, Kladki 24, Poland and the §Département de Biochimie Médicale, Centre Médical Universitaire, Université de Genève, 1, Rue Michel-Servet, 1211 Genève 4, Switzerland

We have used ^{14}C-labeled bacteriophage λO-DNA replication protein as a probe to identify and purify *Escherichia coli* proteases capable of its degradation. In this manner, five different proteases (termed Lop) have been identified capable of degrading λO protein to acid-soluble fragments in an ATP-dependent fashion. One of these activities was purified to homogeneity and shown to be composed of two different polypeptides. The 23,000-Da component (LopP) was identified as the previously characterized ClpP protein, known to complex with ClpA to form the ClpAP, an ATP-dependent protease, capable of degrading casein. The second 46,000-Da component was identified as ClpX (LopC), coded by a gene located in the same operon, but promoter distal to that coding for ClpP (Gottesman, S., Clark, W. P., de Crecy-Lagard, V., and Maurizi, M. R. (1993) *J. Biol. Chem.* 268, 22618–22626). This identification was based on the determination of the sequence of the first 24 amino acid residues of the purified ClpX protein and its identity with that predicted by the DNA sequence. The ClpXP protease is substrate specific, since it degrades casein (known to be degraded by ClpAP), λP, or DnaK proteins slowly or not at all. These results suggest that ClpX protein directs ClpP protease to specific substrates. It is estimated that 50% of all λO-specific protease activity present in crude *E. coli* extracts is due to the ClpXP protease. We propose that transient inhibition of λO degradation observed *in vivo* during the later stages of λ-DNA replication *in vivo* is responsible for the switch from bidirectional to unidirectional replication. One round unidirectional replication will lead to strand separation resulting in a switch from early (theta) to late (sigma) mode of λ-DNA replication.

Energy-dependent proteolysis plays a key role in prokaryotic and eukaryotic cells by controlling the availability of certain short-lived regulatory proteins, ensuring the proper stoichiometry for multi-protein complexes, and helping to eliminate abnormal proteins (1–3). Until now, in *Escherichia coli*, only three ATP-dependent proteases have been purified and characterized (Lon, RecA, and Clp). The Lon protease catalyzes the ATP-dependent degradation of certain denatured proteins and chromogenic peptides, as well as the specific degradation of the bacteriophage λN protein (4–6). Some genetic and biochemical experiments suggest that this protease is also involved in the rapid degradation of the SulA, RcsA, and Tn903 transposase (7–9). The ATP-dependent Lon protease was identified as a heat shock protein, whose synthesis is induced under stress conditions (10). It was also shown that purified RecA protein mediates the cleavage of the bacterial LexA or bacteriophage λcI repressors *in vitro*, and this process is activated by ATP or an analogue of ATP (11, 12).

Recently, the Clp protease of *E. coli* was purified as a two-subunit ATP-dependent protease that contributes to the turnover of abnormal proteins (13, 14). The large ClpA subunit (81,000 Da) has a substrate-stimulated ATPase activity (15, 16). Its amino acid sequence has been conserved to an unusual degree in numerous prokaryotes and eukaryotes (17). Another gene, highly homologous to the *clpA* gene was identified in *E. coli* and named *clpB* (estimated molecular mass of gene product is 84,000 Da) (17). The smaller ClpP subunit (23,000 Da) has low intrinsic peptidase activity which is largely stimulated in the presence of ClpA (15, 18). Neither the ClpA nor ClpP subunit is active alone in degradation of proteins. Recently, both ClpP and ClpB were identified as heat shock proteins (19–21). The *clpA* and *clpP* genes have been cloned and mutations in these genes have been obtained (22). Neither ClpA/ClpP, nor the other well-characterized ATP-dependent protease, Lon, are essential for *E. coli* growth. However, cells lacking ClpA and Lon have a significantly reduced level of abnormal protein turnover (22). Since mutants lacking both ClpA and Lon proteases still show energy-dependent protein degradation, it is clear that other ATP-dependent proteases are present in *E. coli* (22). One of them is a membrane-associated ATP-stimulated protease (23). In addition, the *alp* gene was identified on the basis that it can suppress various *lon* mutant phenotypes, when present in high copy (24). The ability of *E. coli* to rapidly degrade abnormal (as well as some regulatory proteins such as the σ^{32} subunit of RNA polymerase) is inhibited by mutations in any of several heat shock chaperone genes (*dnaK*, *dnaJ*, *grpE*, *groES*, and *groEL*) (25–27). It was proposed that a failure of chaperones to promote protein folding may specifically mark a protein for rapid degradation (28).

In vivo, bacteriophage λ-DNA replication initiates from a unique site (*oriλ*) and proceeds bidirectionally (29). In the first step of this reaction, the bacteriophage λ initiation protein, λO, binds to four direct repeats located in the *oriλ* sequence and forms a DNA-nucleoprotein structure, referred to as "O-some" (30, 31). The λO protein, which triggers the initiation event by introducing the host replication machinery

* This work was supported by Polish Committee for Scientific Research Grant KBN 40001/91/01, National Science Foundation International Exchange Grant INT-8915161, National Institutes of Health grants, and Fonds National Suisse Grant 31-31129-91. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom reprint requests and correspondence should be addressed. Fax/Tel.: 48-58-31-00-72.

at the *oriλ* sequence, is a highly unstable protein. The half-life of λO protein at 37 °C has been estimated to be 1.5 min (32, 33). Until now, λO-specific proteases had not been identified. However, it was shown that expression of the λ gene, *rexB*, prevents degradation of λO protein *in vivo* (34).

In vitro, using a crude protein fraction from an *E. coli* supplemented with purified λO and λP proteins, λ-DNA replication is also bidirectional (35). Similar to the *in vivo* case, λO protein is efficiently degraded (this paper). Surprisingly, reconstitution of the *in vitro* λ-DNA replication system, using only highly purified *E. coli* replication proteins (DnaB, DnaG, DnaK, DnaJ, Ssb, and DNA polymerase III) supplemented with λO and λP proteins, leads to initiation of λ-DNA replication which proceeds predominantly in one direction (from left to right) (36–38).

In this case, λO protein is not degraded (this paper). Thus, we decided to test the hypothesis that, in the purified replication system, the λO protein remains in a complex with *oriλ* DNA and creates a physical barrier for fork movement in the right-to-left direction and that the action of λO-specific proteases triggers the switch from uni- to bidirectional λ-DNA replication *in vivo*.

In this paper we describe the existence of various different factors involved in the degradation of [^{14}C]λO protein to trichloroacetic-soluble counts. One of these enzymatic activities was more extensively investigated and identified as a new regulatory component of the Clp ATP-dependent protease. NH_2-terminal sequencing of the purified protein has identified it as the product of gene *clpX*, the second gene of the *clpP clpX* operon (see accompanying paper, Gottesman *et al.*).

MATERIALS AND METHODS

Bacterial Strains—The *E. coli* W3110 B178 strain has been previously described (35, 40). Plasmid pWPC-9 overexpressing ClpP and ClpX (39) was obtained from Dr. S. Gottesman (National Institutes of Health).

Replication Proteins and in Vitro λ-DNA Replication Assays—The [^{14}C]λO protein was purified as described previously (40). The specific radioactivity of the purified protein was 9,375 cpm[1]/μg. The [^{14}C]λP (5,000 cpm/μg) and DnaK proteins (2,500 cpm/μg) were gifts from Drs. J. Osipiuk and K. Liberek (University of Gdansk), respectively. The [*methyl*-^{14}C]casein was purchased from Du Pont-New England Nuclear. All other replication proteins, λP, DnaB, DnaG, DnaK, DnaJ, GrpE, and Ssb, were purified as described previously (41). Plasmid λ*dv* DNA was also purified as described previously (41). DNA polymerase III was a gift from Dr. J. Kaguni (Michigan State University). The DNA replication assay composed entirely of purified proteins was conducted as described (41). Similarly, the crude enzymatic fraction II of *E. coli* was prepared, and the λ-DNA replication system based on this crude enzymatic fraction was conducted as described previously (42). The λO-DNA complex was isolated during Sepharose 4B chromatography as described previously (40). The serum ClpP and Lon was a generous gift from Dr. S. Gottesman and M. Maurizi.

Protease Assays—The [^{14}C]λO protein (160 ng, 1,500 cpm) was incubated with aliquots (2–10 μl) of protein fractions from the different purification steps of Lop proteases in the presence of 20 mM HEPES/KOH, pH 7.2, 10 mM MgAc, 10 mM ATP, 40 mM creatine phosphate, 0.1 mg/ml creatine kinase, 0.5% Brij 58, and 4 mg/ml bovine serum albumin. The reaction mixture (100 μl), if necessary, was supplemented with ClpP protein (200 ng) and incubated for 30–120 min at 30 °C. The reaction was stopped on ice by the addition of ice-cold trichloroacetic acid (final concentration of 10%). After centrifugation (10 min, 5,000 × *g* at 0 °C), the radioactivity of the trichloroacetic acid-soluble fraction was measured in a toluene/Triton X-100 scintillation fluid. The [^{14}C]casein was used in the protease assay as described (5).

Purification of ClpX—*E. coli* B178 cells were grow in LB medium (15 liters) supplemented with K_2HPO_4 (125 g), KH_2PO_4 (25 g), NH_4Cl (50 g), $MgSO_4$ (20 g), $CaCl_2$ (0.3 g), and sodium citrate (10 g) at 37 °C. During growth, glucose was added to a final concentration of 1%, and the pH was maintained at 7.0 by the addition of NaOH. The bacteria were grown to late exponential phase (OD_{595} = 9), harvested by continuous rotor centrifugation, resuspended in a minimal volume of buffer A (50 mM Tris-HCl pH 7.2 (pH at room temperature), 10% w/v sucrose, 1 mM EDTA), and frozen in liquid nitrogen.

All steps of the purification were carried out at 0–4 °C. A frozen cell suspension of 100 g (wet weight) *E. coli* B178 bacteria was resuspended in 100 ml of buffer A, followed by the addition of 20 ml of buffer B (60 mM Tris-HCl, pH 7.2, 1.2 M NaCl, 1.8 M ammonium sulfate, 0.24 M spermidine-HCl, and 0.012 M dithiothreitol (DTT)). Lysis was initiated by the addition of 0.3 g of lysozyme resuspended in 30 ml of buffer A. The volume of lysate was adjusted to 200 ml by the further addition of buffer A. To ensure complete lysis, after 60 min of incubation at 0 °C the lysate was transferred to 42 °C for 5 min, then returned to ice for an additional 10 min. The lysate was centrifuged in a Beckman R35 rotor at 30,000 revolutions/min at 0 °C for 20 min. To the clear supernatant, streptomycin (20% stock solution, pH 7.0) was added to a final concentration of 5% and precipitation of DNA allowed to proceed during a 30-min incubation at 0 °C. Precipitated DNA was removed by centrifugation (20 min, 30,000 revolutions/min, Beckman R35 rotor at 0 °C). Proteins in the supernatant were precipitated with ammonium sulfate (0.30 g/ml of supernatant). After 30 min at 0 °C, the suspension was centrifuged at 20,000 revolutions/min in a Beckman R35 rotor for 20 min at 0 °C. The pellet was resuspended in 200 ml of buffer G (50 mM Tris-HCl pH 7.2 (pH at room temperature), 5 mM $MgCl_2$, 10 mM DTT, 20% (v/v) glycerol), and directly applied to a P11 phosphocellulose column (5 cm × 15 cm) equilibrated previously with buffer G supplemented with 0.05 M KCl. The P11 column was directly connected to a Q-Sepharose column (2.5 cm × 15 cm) equilibrated with the same buffer. After extensive washing (12 h, 60 ml/h) of the Q-Sepharose column with buffer G (0.05 M KCl), the bound proteins were eluted with a salt gradient (500 ml × 500 ml) from 0.05 to 0.4 M KCl. Fractions active in the λO proteolysis assay were pooled and diluted 4-fold with buffer G and applied to a Bleu M Tris-acryl (IBF Biotechnics) column (2.5 cm × 7 cm) previously equilibrated with buffer G (0.05 M KCl). The column was washed with the same buffer for 12 h, 40 ml/h. λO proteolytic activities were eluted from the column by a salt gradient (250 ml × 250 ml) from 0.05 to 0.6 M KCl. Active fractions were directly applied to a hydroxylapatite column (1.5 cm × 2.5 cm) previously equilibrated with buffer G (0.05 M KCl). The column was washed with the same buffer for 1 h, 20 ml/h. The bound protein fractions were eluted with a gradient (75 ml × 75 ml) from 0 to 0.2 M $KHPO_4$, pH 6.9. To active fractions, 10% Triton X-114 was added to a final concentration of 0.015% and dialysis was carried out for 2 × 4 h against 0.5 liter of buffer G supplemented with 0.015% of Triton X-114, changing to fresh buffer after the first 4-h period. The dialyzed fractions were applied to a heparin-agarose column (1 cm × 8 cm) equilibrated with buffer G. After washing the column for 16 h, 8 ml/h, ClpX protein (>90% pure) was eluted with a gradient (30 ml × 30 ml) from 0 to 0.2 M KCl.

Purification of ClpP Protein—B178 bacteria carrying plasmid pWPC-9 were grown at 37 °C in LB medium to OD_{595} = 1.2. The frozen cell suspension was thawed, lysed, and cytosolic proteins precipitated with ammonium sulfate as described in the case of ClpX purification. The first two columns, P11 and Q-Sepharose, were run exactly as described for the purification of ClpX protein, ClpP activity being separated from ClpX activity on the Bleu M Tris-acryl column. ClpX activity was eluted from the Bleu M Tris-acryl column in the range of 0.1 to 0.2 M KCl. ClpP protein was eluted between 0.3 and 0.4 M KCl. The active ClpP protein fractions were applied to a hydroxylapatite column and heparin column under the conditions described for ClpX purification. ClpP activity was eluted from the hydroxylapatite column between 30 and 50 mM $KHPO_4$ (using a gradient from 0 to 80 mM KHPO4). It was eluted from the heparin column between 100 and 170 mM KCl (using a gradient from 0 to 250 mM KCl). After the heparin column, ClpP protein was more than 90% pure and could be stored at 4 °C for more than 6 months without substantial loss of activity. The NH_2-terminal domain of the 23,000-Da protein was sequenced and shown to be identical to the amino acid sequence predicted from the *clpP* gene with the exception that it is missing a 14 amino acid precursor peptide from the amino terminus (results not shown).

[1] The abbreviations used are: cpm, counts/min; DTT, dithiothreitol.

RESULTS

Monitoring of λO Protein during DNA Synthesis in the Crude and Purified in Vitro λ-DNA Replication Systems—The ^{14}C-labeled λO protein, a key λ replication protein which specifically binds the *ori*λ sequence and, with the help of λP, attracts host replication machinery to this site (41), was purified to homogeneity from *E. coli* bacteria carrying the pMY17-3 plasmid (41) grown in minimal medium containing [^{14}C]protein hydrolysate. As published previously, purified [^{14}C]λO forms a stable complex with DNA containing the *ori*λ sequence. This complex appears in the excluded volume following chromatography on a Sepharose 4B column (Fig. 1; see also Ref. 40). The *ori*λ-λO complex is specific because substitution of the *ori*λ sequence by *ori*82 severely reduces the binding of λO to DNA (40). Purified [^{14}C]λO protein was fully active in both crude and purified λ-DNA replication systems (results not shown), so we monitored the formation of the [^{14}C]λO-DNA complex during these events (Fig. 1). It is clear that after initiation of λ-DNA replication in our purified system, most of the ^{14}C-labeled λO protein (80–90%) remains bound to λ-DNA (Fig. 1). The rest of the [^{14}C]λO protein (10–20%) is released during the initiation process and behaves as a protein aggregate on a Sepharose 4B column (Fig. 1). Surprisingly, when λ DNA synthesis was conducted in the crude replication system, most of the radioactivity appeared in the included volume of the Sepharose 4B column. However, following SDS-polyacrylamide electrophoresis of these fractions, we were unable to detect the presence of any radioactive λO protein (results not shown). This result suggests that during initiation of λ-DNA replication, using the crude enzymatic fraction, the majority of the λO protein was degraded to small polypeptide fragments or single amino acids.

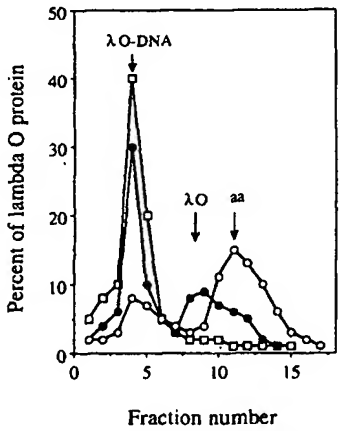


FIG. 1. **Binding of λO protein to λ-DNA in a crude and purified *in vitro* λ-DNA replication system.** The purified ^{14}C-labeled λO protein (1 μg) was incubated in 120 μl of reaction with 1.5 μg of λ*dv* DNA alone (□); in the presence of crude enzymatic fraction under conditions leading to λ-DNA replication (○); in the presence of purified replication proteins under condition leading to GrpE-dependent λ-DNA replication (●). After 12 min of incubation at 30 °C, 100 μl of the reaction was loaded directly on a Sepharose 4B column, whereas 20 μl of [^{3}H]TTP (1 μl of 1 mCi/ml) was added to the rest, and the reaction was allowed to proceed for an additional 20 min. The reaction was stopped and the amount of [^{3}H]TMP incorporated into DNA was estimated as described previously (40). The conditions for the crude enzymatic replication system were as described in Ref. 42 with the exception that 500 μg of Fraction II (isolated from B178 bacteria) was used; the level of incorporation of [^{3}H]TMP to DNA in the control experiment was estimated to be 220 pmol of dNMP. The conditions for the purified replication system were those described in Ref. 41. The level of incorporation of [^{3}H]TMP to DNA in the control experiment was 180 pmol of dNMP. The chromatography on Sepharose 4B was carried out as described in Ref. 40. Fractions (100 μl) were collected and the ^{14}C counts derived from ^{14}C-labeled λO were estimated in a scintillation counter.

λO Proteolytic Activity in the Crude Protein Extract—When [^{14}C]λP, -λO, and -DnaK proteins were incubated at 30 °C with the crude protein fraction isolated from *E. coli* under conditions known to support λ-DNA replication, only [^{14}C]λO protein was efficiently degraded to small peptides or single amino acids (Fig. 2*A*). This result suggests the existence of a protease(s) which specifically recognizes λO protein. Such a protease(s) must be very processive since we were unable to detect any proteolytic intermediates even following separation on a 20% polyacrylamide gel (result not shown). After optimization of the proteolytic assay conditions (see "Materials and Methods"), it became clear that λO protein degradation is dependent on the presence of ATP. This reaction, however, is independent of the presence of λ-DNA with which λO protein specifically interacts (Fig. 2*B*). The results shown in Table I indicate that serine protease inhibitors partially block λO degradation (see also "Discussion").

Identification of λO-specific Proteolytic Activities—To identify the λO-specific proteolytic activities, wild type *E. coli* B178 bacteria were lysed and the soluble proteins precipitated with ammonium sulfate, as described under "Materials and Methods." The precipitated proteins were applied to a Q-Sepharose column. Most of the λO proteolytic activity (90%) bound to the column and could be quantitatively eluted with a KCl gradient as a single peak (~165 mM KCl) (Fig. 3). Until now, most of the ATP-dependent proteases from *E. coli* had been isolated using casein as a substrate. Thus, we tested the various fractions for [^{14}C]casein degradation activity. Since the overlap between λO protease activity and casein protease activity is not perfect, it is clear that at least some proteases which degrade casein do not efficiently degrade λO protein (Fig. 3).

Because of previously documented examples of multisubunit proteases, *e.g.* ClpAP (15, 18), we tested whether the proteolytic activity found in the single peak at 165 mM KCl could be enhanced by the addition of other fractions. This was done by taking 1/10 the amount normally used for the

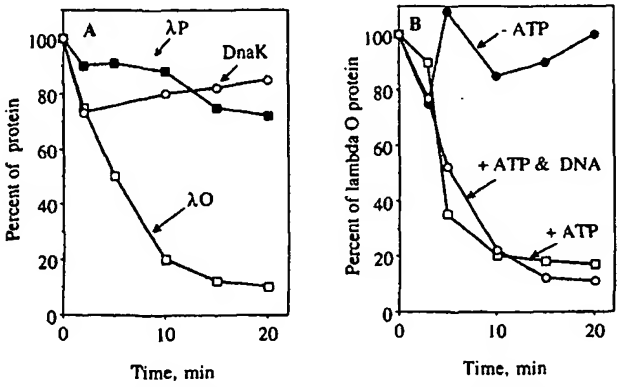


FIG. 2. **λO proteolytic activity in crude enzymatic fraction derived from B178 *E. coli* bacteria.** *A*, the purified ^{14}C-labeled λO protein (200 ng) (□), DnaK protein (200 ng) (○), or λP protein (200 ng) (■) was incubated in 25-μl assays at 30 °C for the desired time under the exact conditions leading to λ*dv* DNA replication as described in Ref. 42, except that the dNTP mixture was omitted and 100 μg of B178 fraction II was used. The reaction was stopped by the addition of 200 μl of 10% trichloroacetic acid, and the amount of trichloroacetic acid-soluble counts was estimated as described under "Materials and Methods." *B*, the purified ^{14}C-labeled λO protein was incubated at 30 °C in the replication assay as described above (○). In the control experiment, ATP (●) or λ DNA (□) was omitted.

TABLE I

Effect of various serine protease inhibitors on ATP-dependent degradation of [^{14}C]λO protein by crude E. coli enzymatic fraction and purified ClpXP protease

The 160 ng of [^{14}C]λO protein was incubated with 100 μg of fraction II or 8.7 pmol of ClpP with 1.3 pmol of ClpX in the presence or absence of serine inhibitors for 60 min as described under "Materials and Methods." PMSF, phenylmethylsulfonyl fluoride; TPCK, L-1-tosylamido-2-phenylethyl chloromethyl ketone; TLCK, 1-chloro-3-tosylamido-7-amino-2-heptanone.

Addition	Relative activity	
	Crude enzymatic fraction	ClpXP
mM	%	
None	100	100
PMSF		
0.25	76	98
5	55	89
10	31	88
TLCK		
0.25	42	92
0.5	33	85
1	25	71
2.5	19	23
TPCK		
0.25	92	96
0.5	79	95
1	71	97
2.5	55	93

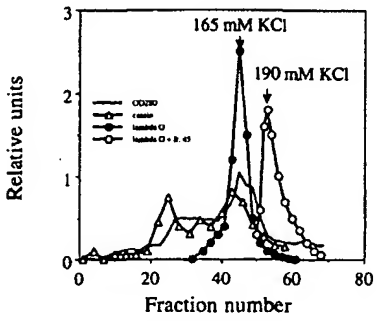


FIG. 3. **Separation of λO proteases on a Q-Sepharose column.** Bacteria were lysed, and the protein extract was precipitated with ammonium sulfate as described under "Materials and Methods." The protein extract was dialyzed overnight against buffer G and applied on a Q-Sepharose column (2.5 × 15 cm). The adsorbed proteins were eluted with a salt gradient (0.05–0.4 M). The absorption of each fraction (....) was measured in a spectrophotometer (one relative unit represents A_{280} = 1). The activity of λO-specific proteases (●) was estimated by incubating 10 μl of each particular fraction with ^{14}C-labeled λO protein as described under "Materials and Methods." One relative unit represents 300 trichloroacetic acid-soluble counts/min released from [^{14}C]λO protein during a standard proteolysis assay carried out for 30 min at 35 °C. One μl of fraction 45 was incubated with 10 μl of material taken from fractions 49 to 70 under the same conditions as those described for the standard λO proteolysis assay (○). The ability to hydrolyze ^{14}C-labeled casein (△) was tested as described in Ref. 5 (one unit represents 300 trichloroacetic acid-soluble counts/min released during hydrolysis of ^{14}C-labeled casein).

assay (material from fraction 45 at this concentration gives only a weak response in the λO proteolytic assay) and complementing this with fractions collected from the Q-Sepharose column at different concentrations of KCl. As shown in Fig. 3, a second well defined peak of λO proteolytic activity was observed. This last result suggests that at least some of the λO-specific proteases possess a multicomponent structure and that these components are partially separated on the Q-Sepharose column. When the same assay was done in reverse, that is, supplementing a low concentration of the material from the 190 mM KCl peak with material from the 165 mM KCl peak, a similar stimulation of λO protease activity was observed (results not shown). Almost no detectable λO proteolytic activity was detected when ATP was omitted from our standard proteolytic assay (results not shown), suggesting that most of the λO-specific proteases are ATP-dependent.

Because the situation described above is strongly reminiscent of the behavior of the ClpP subunit during chromatography on Q-Sepharose (15), the material from the 190 mM KCl peak was further purified using Bleu M Tris-acryl, hydroxylapatite, and heparin-agarose columns as described under "Materials and Methods." During the purification, we monitored λO protease activity using our standard proteolytic assay supplemented, as before, with a low concentration of material from the 165 mM KCl peak of the Q-Sepharose column. Using this protocol, we isolated a single polypeptide of 23,000 Da, which was identified as the product of the *clpP* gene by Western blot analysis and sequencing of the NH_2-terminal domain (results not shown).

The λO proteolytic activities could be separated using a combination of P11 phosphocellulose, heparin-agarose, and hydroxylapatite column chromatography (Fig. 4). Using this approach, we detected at least five different factors capable of hydrolyzing λO. We call these activities "Lop" (lambda O proteases). Only the LopA activity does not need to be complemented by other factors. After extensive purification LopA protein was found to be identical (Western blot and NH_2-terminal sequence analysis) with Lon protein.[2] The LopR activity is not complemented by ClpP (LopP) but is highly stimulated by the presence of ClpX (LopC) or LopD. At this stage we cannot exclude the possibility that LopR is not a new proteolitic component but instead represents unprocessed ClpP or ClpP complexed with other factor(s). Two peaks of ClpP activity were also observed previously (39). The (ClpX) LopC and LopD activities were also complemented by ClpP (LopP). Below, we characterize more extensively the LopC factor (ClpX). Our recent experiment suggest, that LopD may be identical with ClpA.[3]

Purification of the ClpX Protein—Using the protocol outlined under "Materials and Methods," we purified to near homogeneity, starting with wild type *E. coli* B178 cells, the ClpX protein (Table II). To monitor the activity of this protein, we followed its ability to help degrade [^{14}C]λO protein in the assay supplemented with purified ClpP protein.

The P11 phosphocellulose column was used before Q-Sepharose to eliminate both Lon and ClpA activities. The Lon and ClpA proteins bind to the P11 column whereas the ClpX activity elutes in the void volume.[4] ClpX protein was applied directly to the Q-Sepharose column, followed by purification on Bleu M Tris-acryl, hydroxylapatite, and heparin columns. The ClpX activity was concentrated on an additional hydroxylapatite column. The fractions from the heparin column were analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 5). There is a perfect correlation between λO protease activity and the presence of an approximately 46,000-Da protein. A silver-stained gel shows the different steps of purification of the ClpX protein (Fig. 6). A polypeptide of approximately 46,000-Da molecular mass, some contaminating ClpP protein (23,000 Da), and an approximately 200-kDa protein were

[2] D. Wojtkowiak, C. Georgopoulos, and M. Zylicz, unpublished results.

[3] D. Wojtkowiak, unpublished results.

[4] This is true only for non-overproducing ClpX extracts.

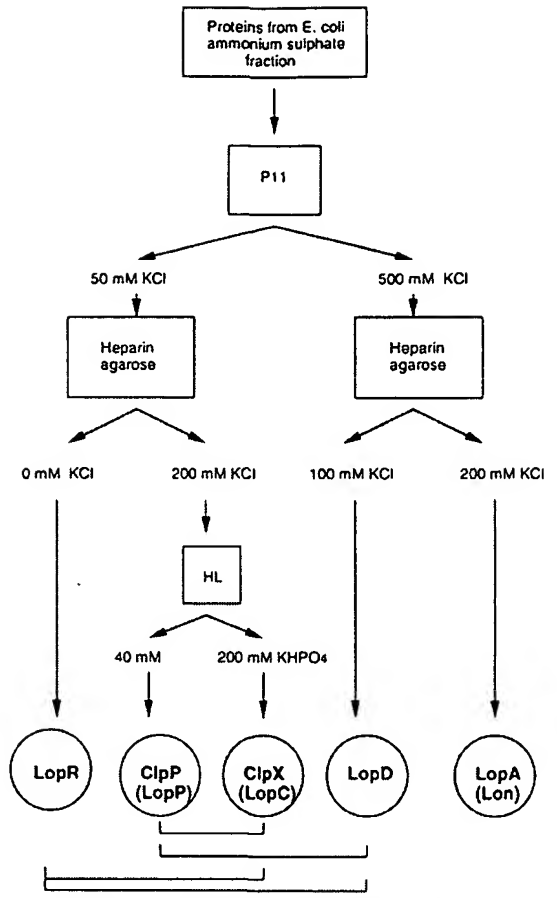


FIG. 4. **Isolation of different factors involved in λO degradation.** *E. coli* B178 bacteria (100 g wet weight) were lysed and proteins precipitated with ammonium sulfate and applied to a phosphocellulose (P11) column as described under "Materials and Methods." Proteins found in the void volume were applied directly to a heparin column under the conditions described (see "Materials and Methods"). λO proteolytic activity found in the void volume of the heparin column was named LopR. LopR requires the addition of ClpX or LopD for efficient λO degradation. Protein which eluted from the heparin column in the presence of 200 mM KCl represents two overlapping activities which could be separated during hydroxylapatite column chromatography. These were determined to be the ClpP and ClpX activities which elute at 40 and 200 mM $KHPO_4$, respectively. Proteins which bound to the phosphocellulose P11 column and eluted in the presence of 500 mM KCl were dialyzed in buffer G (see "Materials and Methods") and applied to a heparin column. LopD activity (which requires LopR or ClpP for efficient λO degradation) elutes in the presence of 100 mM KCl. LopA activity (which does not require the addition of other factors for λO degradation) elutes in the presence of 200 mM KCl. Further purification and analysis showed that LopA is actually the previously identified Lon protease.

TABLE II

Purification of ClpX protein from E. coli wild type bacteria

Purification was carried out as described under "Materials and Methods." One unit of ClpX protease activity was defined as 0.01 pmol of [^{14}C]λO protein hydrolyzed/min in our standard assay in the presence of saturating amounts of ClpP protein.

Purification step	Protein	Volume	Activity	Specific activity
	mg	*ml*	*units*	*unit/μg*
1. Ammonium sulfate	450	10	190,000[a]	0.42
2. Q-Sepharose	62	125	175,000	2.8
3. Bleu M Tris-acryl	6.1	82	82,000	13.3
4. Hydroxylapatite I	1.5	24	38,000	26.6
5. Heparin-agarose	0.045	18	18,000	400
6. Hydroxylapatite II	0.022	4.5	9,500	420

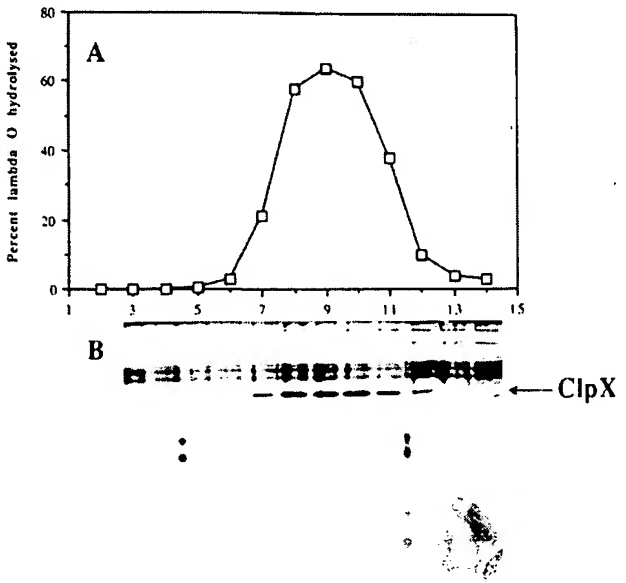


FIG. 5. **Purification of ClpX component on a heparin-agarose column.** *E. coli* bacteria harboring plasmid pWPC-9, overproducing ClpP and ClpX proteins, were grown at 37 °C in the presence of ampicillin to OD_{600} = 1.0. The cells were lysed and ClpX protein purification was carried out up to the heparin-agarose column as described under "Materials and Methods." *A*, λO proteolytic activity of protein fractions eluted from the heparin-agarose column by increasing concentrations of KCl (see "Materials and Methods" for details). An aliquot (8 μl) of each fraction was analyzed in our standard λO proteolytic assay in the presence of 8.7 pmol of ClpP protein. The assay was conducted for 20 min at 30 °C and trichloroacetic acid-soluble counts, derived from the degradation of [^{14}C]λO protein, were estimated as described under "Materials and Methods." *B*, SDS-polyacrylamide (12.5%) gel electrophoresis of fractions from the heparin column for which the λO proteolytic activity was measured. The gel was silver-stained.

observed. We sequenced the NH_2 terminus of the 46,000-Da polypeptide and showed that the sequence is identical to the amino acid sequence predicted for a gene located between the *clpP* and *lon* genes (39) (Table III; 10 min on the *E. coli* map; see accompanying paper).

Stability of ClpX Activity—The activity of the purified ClpX protein is extremely unstable when stored without the ClpP component or at low salt concentrations (<50 mM KCl). The presence of 20–50% (v/v) glycerol, 10 mM DTT, and 0.015% Triton X-114 substantially stabilizes ClpX activity. The ClpX protein could not be frozen and stored at −70 °C without substantial loss of activity. When stored at −15 °C in 20% glycerol, 70% of the activity is lost during a 1-month period. The ClpX protein has a strong tendency to aggregate and bind to glass and dialysis tubing. The addition of 0.015% Triton X-114 helps prevent this aggregation.

Properties of the Purified ClpXP Protease—Degradation of [^{14}C]λO protein by the ClpXP protease was absolutely dependent on the presence of ATP (Fig. 7). In agreement with this, the presence of an ATP regenerating system stimulated degradation of the λO protein (Table IV). Of the various nucleotides tested, only dATP could partially substitute for ATP in this reaction. Neither component alone possessed any λO-specific proteolytic activity (Fig. 8). The kinetics of λO degradation was linear up to 60 min after initiation of the reaction (Fig. 9). In a control experiment, we could not detect, under our assay conditions, any degradation of [^{14}C]casein,

suggesting that purified ClpX component directs ClpP protease to specific substrates (Fig. 9). The presence of DTT at a high concentration inhibited the proteolysis reaction (Table V), but the presence of 5–10 mM DTT not only stabilized the protease but also stimulated the degradation reaction (Table V). The presence of KCl, $ZnSO_4$, glutathione, *N*-ethylmaleimide, *p*-chloromercuribenzoate, iodoacetamide, and EDTA substantially inhibited the ClpX-dependent λO protease (Table V).

DISCUSSION

It was shown previously that the λO replication protein, which is required for the initiation of λ-DNA replication, is extremely susceptible to ATP-dependent proteolysis (32, 33, 44, 45). Until now, most of the known ATP-dependent proteases in *E. coli* (Lon and ClpAP) were identified using casein as a substrate for degradation. Our use of the more biologically relevant bacteriophage λ-DNA replication protein, λO, as a substrate has led to the identification of at least five different ATP-dependent factors involved in λO proteolysis. In this paper we characterized one of these, ClpX (defined as LopC in our original nomenclature). In the presence of ClpP, ClpX specifically degrades λO protein and not casein. We estimate that at least 50% of the λO-dependent activity present in a crude *E. coli* fraction is due to ClpXP activity (results not shown). The other factors that we identified degrade λO protein less efficiently. In agreement with this conclusion, Gottesman *et al.* (accompanying paper) have shown that mutations in either *clpP* or *clpX* substantially extend the half-life of λO protein *in vivo*.

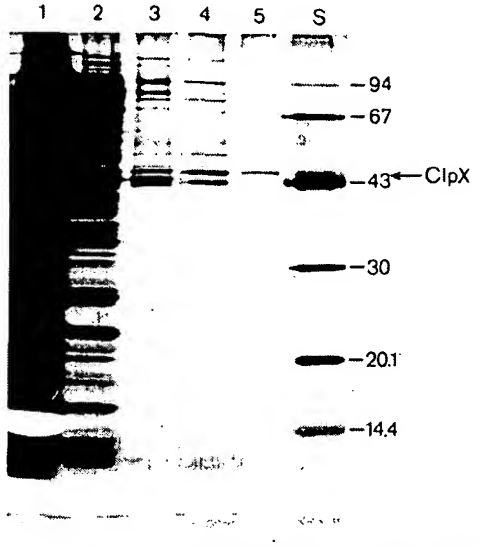


FIG. 6. **SDS-polyacrylamide gel electrophoresis of fractions taken from different steps of ClpX purification.** The ClpX protein was purified from wild type *E. coli* B178 bacteria as described under "Materials and Methods." Aliquots containing equal amounts of λO proteolytic activity (45 units, see Table II for definition of units) from each step of the purification were resolved by 12.5% SDS-polyacrylamide gel electrophoresis, followed by silver staining. *Lane 1*, ammonium sulfate fraction; *lane 2*, active fractions pooled from Q-Sepharose column; *lane 3*, active fractions pooled from the Bleu M Tris-acryl-agarose column; *lane 4*, active fractions pooled from the hydroxyapatite column; *lane 5*, fraction 9 from the heparin-agarose column; *lane S*, molecular weight standards.

It was shown previously that ClpA, which contains an ATP-binding site, in combination with ClpP, which contains a serine peptidase active site, promotes degradation of casein (13–16). The ClpX component appears also to be a specificity component for ClpP. For example, ClpX, but not ClpP, specifically binds to a λO-agarose column and is eluted at a concentration of 0.5–1 M KCl.[2] Our results are in agreement with the proposal that in *E. coli*, the ClpP component plays the role of a "master" protease which is "attracted" to different substrates by different specificity factors (Fig. 10). It is not clear if these proteins form a multienzymatic structure similar to that of the ATP/ubiquitin-dependent protease system isolated from eukaryotic cells (46). It could be that the newly discovered ClpB polypeptide, whose sequence is homologous to that of ClpA, can also complex with ClpP to give rise to yet another ATP-dependent protease (47).

Substituting the ClpX component for the ClpA subunit in the ClpP-dependent reaction does not change its overall qualitative behavior in the presence of different inhibitors and nucleotides, *e.g.* cysteine inhibitors (*N*-ethylmaleimide, *p*-chloromercuribenzoate, glutathione, iodoacetamide, and Zn^{2+}) efficiently block λO degradation. However, two quantitative differences do exist, *i.e.* the ClpXP enzyme can be inhibited by 6-fold lower levels of *N*-ethylmaleimide than the ClpAP protease (48). At the same time, *p*-chloromercuribenzoate (0.1 mM) inhibits ClpXP protease activity 40-fold less efficiently than the ClpAP enzyme (48). Purified ClpXP is not efficiently inhibited by phenylmethylsulfonyl fluoride and L-1-tosylamido-2-phenylethyl chloromethyl fluoride serine protease inhibitors; only 1-chloro-3-tosylamido-7-amino-2-heptanone inhibits efficiently the degradation of λO protein catalyzed by purified ClpXP (Table I). There is no perfect correlation between the ability of serine protease inhibitors to block the λO degradation by crude enzymatic fraction and purified ClpXP protease. This is due to the fact that ClpX activity represents only about 50% of total proteolytic activity toward λO present in crude enzymatic fraction. In addition, we cannot exclude the possibility that the presence of other factors, like

TABLE III

NH_2-terminal sequence of the ClpX protein

Protein sequencing was performed as described in Ref. 49. The ClpX protein was resolved by SDS-polyacrylamide gel electrophoresis. Following electrophoresis, the protein was transferred from the gel to polyvinylidine difluoride membrane (Immobilon; Millipore Corp.). The membrane fragment containing the 46,000-Da protein was used for sequencing in a model 477A sequencer (Applied Biosystems). The analysis was performed by Dr. Robert Schackmann at the University of Utah Biotechnology Center.

```
                                        nt 1130
                                          |
(1) from DNA sequence (39)          CCC ATG ACA GAT AAA CGC AAA GAT GGC TCA GGC
(2) predicted aa sequence             MET-THR-ASP-LYS-ARG-LYS-ASP-GLY-SER-GLY-
(3) aa sequence determined               THR-ASP-LYS-ARG-LYS-ASP-GLY-SER-GLY-
                                                    nt 1200
                                                      |
AAA TTG CTG TAT TGC TCT TTT TGC GGC AAA AGC CAG CAT GAA GTG    (1)
LYS-LEU-LEU-TYR-CYS-SER-PHE-CYS-GLY-LYS-SER-GLN-HIS-GLU-VAL    (2)
LYS-LEU-LEU-TYR-    -SER-PHE-    -GLY-LYS-SER-GLN-HIS-GLU-VAL    (3)
```

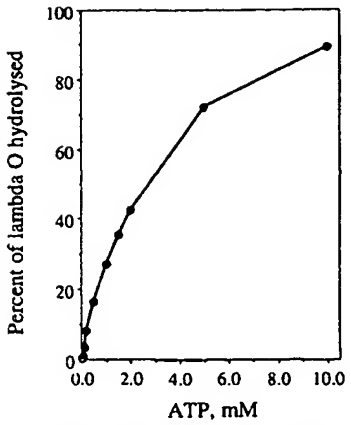


FIG. 7. **ATP dependence of λO degradation catalyzed by the ClpXP protease.** The λO proteolysis assay was carried out for 80 min at 30 °C in the presence of 8.7 and 1.3 pmol of ClpP and ClpX proteins, respectively, under the conditions described under "Materials and Methods."

TABLE IV

Nucleotide specificity of the ClpXP protease

The indicated nucleotides were added to a final concentration of 5 mM in the standard assay mising CP and CK (see "Materials and Methods") and incubated for 60 min at 30 °C. The amount of trichloroacetic acid-soluble [^{14}C]λO protein was determined. CP, creatine phosphate; CK, creatine kinase.

Nucleotide	% of λO protein degradation
None	0
ATP + CP/CK	100
ATP	68
GTP	0
CTP	0
UTP	0
dATP	34
dCTP	0
dGTP	0
dTTP	0
AMP	0
ADP	0

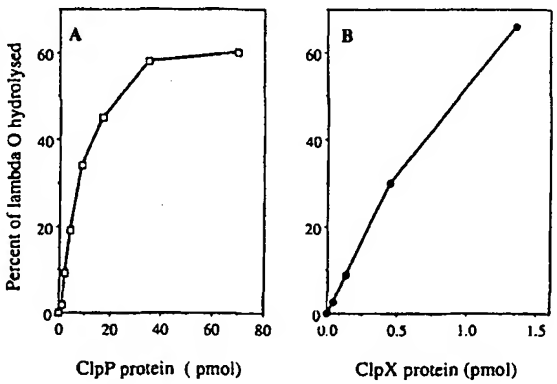


FIG. 8. **Titration of ClpP and ClpX proteins in the λO proteolysis assay.** ^{14}C-λO degradation was analyzed after incubation of λO protein with ClpP and ClpX proteins for 15 min at 30 °C as described under "Materials and Methods." *A*, increasing amounts of ClpP protein in the presence of 0.42 pmol of ClpX. B. Increasing amounts of ClpX protein in the presence of 8.7 pmol of ClpP protein.

Lon or ClpA can stimulate ClpXP-dependent degradation of λO protein.

The *clpX* gene is located immediately downstream of *clpP*,

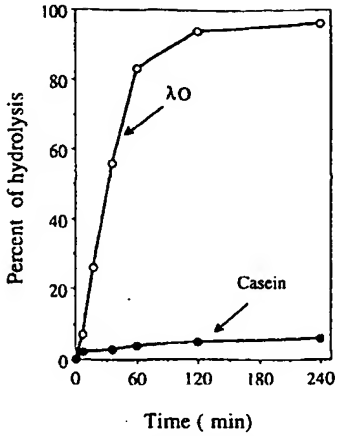


FIG. 9. **Kinetics of λO degradation catalyzed by ClpXP protease.** [^{14}C]λO protein (320 ng) was incubated for the indicated times at 30 °C in the presence of 0.22 and 17 pmol of ClpX and ClpP proteins, respectively, under the buffer conditions described under "Materials and Methods" (○). The reactions were stopped by the addition of trichloroacetic acid and soluble counts were measured in a scintillation counter as described under "Materials and Methods." In the control experiments, 1 μg of [^{14}C]casein was incubated with 17 and 0.22 pmol of ClpP and ClpX, respectively (●).

TABLE V

Effect of various protease inhibitors on ClpXP-dependent degradation of ^{14}C-labeled λO

The 160 ng of [^{14}C]λO protein was incubated with 8.7 pmol of ClpP and 1.3 pmol of ClpX (in the presence or absence of additional chemicals) for 60 min as described under "Materials and Methods." NEM, *N*-ethylmaleimide; PMB, *p*-chloromercuribenzoate.

Addition	Relative activity
mM	%
None	100
Iodoacetamide	
10	78
20	44
50	9
Dithiothreitol	
5	121
20	97
80	28
Glutathione	
8	29
KCl	
100	100
200	6.6
300	0.4
NEM	
0.1	34
1	7.5
PMB	
0.1	43
0.2	19
$ZnSO_4$	
0.1	66
1	4.5

the two genes forming an operon (see accompanying paper). Since ClpP has been shown to be a heat shock protein, under σ^{32}-dependent regulation (21), it is most likely that ClpX is also a heat shock protein (see accompanying paper). Interest-

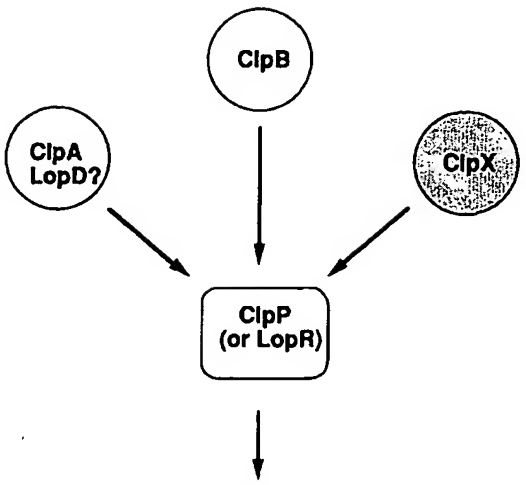


FIG. 10. **The ClpP protease subunit can be coupled to many different factors.** ATP-dependent proteolysis of different protein substrates could be achieved in *E. coli* bacteria by coupling different specificity factors to the ClpP component (see "Discussion" for details).

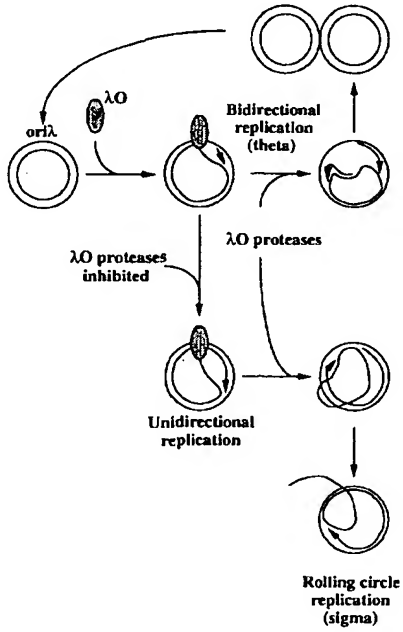


FIG. 11. **Hypothetical model of the involvement of a λO-specific protease in switching from bidirectional to unidirectional theta λ-DNA replication and consequently to the rolling circle mode of DNA replication.** See "Discussion" for details.

ingly, the ClpA subunit in the ClpAP protease is not under heat shock regulation (39). Thus, it could be that the ClpXP protease evolved to deal with the increased need for proteolysis caused by increased denaturation or aggregation, under heat shock conditions. In this respect, it is interesting that Lon (the LopA protease found in our system; result not shown) is also under σ^{32}-dependent regulation (10).

A possibility that cannot be excluded by our data is that ClpX, unlike ClpA, is not a *bona fide* subunit of ClpP. Rather, ClpX binds the λO substrate protein and "presents" it to the ClpP protease in a form that can be recognized and readily hydrolyzed by ClpP. In this model the ClpX protein plays the role of a "chaperone" for λO. In essence, this is the role that ClpA could also play in the ClpAP protease complex, except that the two subunits stably interact.

The physiological role of ClpXP protease in λO degradation is still not clear. *In vivo* as well as *in vitro*, using crude enzymatic fractions purified from *E. coli*, λ-DNA replication proceeds bidirectionally (29). In contrast, reconstituted λ-DNA replication using highly purified proteins is unidirectional and proceeds from left to right (36–38). It was suggested previously that λO protein, which remains bound to the *ori*λ sequence in the purified system, creates a physical barrier that impedes DnaB helicase from unwinding λ-DNA in the right to left direction from *ori*λ (36). Unidirectional replication, after one round, could lead to strand separation, resulting in a switch from the early (theta) to the late (sigma) mode of λ-DNA replication (36). From *in vivo* observations (32, 33), as well as our observations in the crude and purified replication systems, we propose that λO protease(s) could be a factor triggering the switch from uni- to bidirectional replication, and that transient inhibition of this activity could be involved in the switch from early to late modes of λ-DNA replication (Fig. 11).

Acknowledgments—We thank Drs. S. Gottesman and M. Maurizi for sharing unpublished data with us, especially for pointing out to us the fact that the NH_2-terminal amino acid sequence of ClpX perfectly matches that predicted by the *orf* present downstream of *clpP*. We also thank Dr. D. B. Ang for a critical reading of the manuscript.

REFERENCES

1. Bond, J. S., and Butler, P. E. (1987) *Annu. Rev. Biochem.* **56**, 333–364
2. Gottesman, S. (1989) *Annu. Rev. Genet.* **23**, 163–198
3. Miller, C. G. (1987) in *Escherichia coli & Salmonella typhimurium: Cellular & Molecular Biology* (Neidhardt, F. C., Ingraham, J., Low, K. B., Magasanik, B., Schaechter, M., and Umbarger, H. E., eds) pp. 680–691, American Society for Microbiology, Washington, D. C.
4. Charette, M., Henderson, G., W., and Markovitz, A. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 4728–4732
5. Chung, C. H., and Goldberg, A. L. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 4931–4935
6. Maurizi, M. R. (1987) *J. Biol. Chem.* **262**, 2696–2703
7. Mizusawa, S., and Gottesman, S. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 358–362
8. Torres-Cabassa, A. S., and Gottesman, S. (1987) *J. Bacteriol.* **169**, 981–989
9. Derbyshire, C., Kramer, M., and Grindley, N. D. F. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4048–4052
10. Phillips, T. A., VanBogelen, R. A., and Neidhardt, F. C. (1984) *J. Bacteriol.* **159**, 283–287
11. Craig, N. L., and Roberts, J. W. (1980) *Nature* **83**, 26–29
12. Craig, N. L., and Roberts, J. W. (1981) *J. Biol. Chem.* **256**, 8039–8044
13. Katayama, Y., Gottesman, S., Pumphrey, J., Rudikoff, S., Clark, W. P., and Maurizi, M. R. (1988) *J. Biol. Chem.* **263**, 15226–15236
14. Katayama-Fujimura, Y., Gottesman, S., and Maurizi, M. R. (1987) *J. Biol. Chem.* **262**, 4477–4485
15. Hwang, B. J., Woo, K. M. Goldberg, A., L., and Chung, C. H. (1988) *J. Biol. Chem.* **263**, 8727–8734
16. Gottesman, S., and Maurizi, R. M. (1992) *Microbiol. Rev.* **56**, 592–621
17. Gottesman, S., Squires, C., Pichersky, E., Carrington, M., Hobbs, M., Mattick, J. S., Dalrymple, B., Kuramitsu, H., Shiroza, T., Foster, T., Clark, W. P., Ross, B., Squires, C. L., and Maurizi, M. R. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 3513–3517
18. Woo, K. M., Chung, W. J., Ha, D. B., Goldberg, A. L., and Chang, C. H. (1989) *J. Biol. Chem.* **264**, 2088–2091
19. Kitagawa, M., Wada, C., Yoshioko, S., and Yura, T. (1991) *J. Bacteriol.* **173**, 4247–4253
20. Squires, C., Pedersen, S., Ross, B. M., and Squires, C. (1991) *J. Bacteriol.* **173**, 4254–4262
21. Kroh, H. E., and Simon, L. D., (1990) *J. Bacteriol.* **172**, 6026–6034
22. Katayama, Y., Gottesman, S., Pumphrey, J., Rudikof, S., Clark, W. P., and Maurizi, M. R. (1988) *J. Biol. Chem.* **263**, 15226–15236
23. Voellmy, R. W., and Goldberg, A. L. (1981) *Nature* **290**, 419–421
24. Trempy, J. E., and Gottesman, S. (1989) *J. Bacteriol.* **171**, 3348–3353
25. Tilly, K., Spence, J., and Georgopoulos, C. (1989) *J. Bacteriol.* **171**, 1585–1589
26. Straus, D. B., Walter, W. A., and Gross, C. A. (1988) *Genes & Dev.* **2**, 1851–1858
27. Keller, J. A., and Simon, L. D. (1988) *Mol. Microbiol.* **2**, 31–41
28. Sherman, M., and Goldberg, A. L. (1992) *EMBO J.* **11**, 71–77
29. Furth, M. E., and Wickner, S. H. (1983) in *Lambda II* (Hendrix, R. W., Roberts, J. W., Stahl, F. W., and Weisberg, R. A., eds) pp. 145–173, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
30. Schnös, M., Zahn, K., Inman, R. B., and Blatner, F. R. (1988) *Cell* **52**, 385–395
31. Dodson, M., Roberts, J., McMacken, R., and Echols, H. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 4678–4682
32. Lipinska, B., Podhajska, A., and Taylor, K. (1980) *Biochem. Biophys. Res. Commun.* **92**, 120–126

33. Gottesman, S., Gottesman, M., Show J., and Pearson, M. L. (1981) *Cell* **24**, 225–233
34. Schoulaker-Schwarz, R., Dekel-Gorodetsky, L., and Engelberg-Kulka, H. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 4996–5000
35. Wold, M. S., Mallory, J. B., Roberts, J. D., LeBowitz, J. H., and McMacken, R. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 6176–6180
36. Dodson, M., Echols, H., Wickner, S., Alfano, C., Mensa-Wilmot. K., Gomes, B., LeBowitz, J. H., Roberts, J. D., and McMacken R. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 7638–7642
37. Mensa-Wilmot, K., Seaby, R., Alfano, C., Wold, M. S., Gomes, B., and McMacken, R. (1989) *J. Biol. Chem.* **264**, 2853–2861
38. Zylicz, M., Ang, D., Liberek, K., and Georgopoulos, C. (1989) *EMBO J.* **8**, 1601–1608
39. Maurizi, M. R., Clark, W., P., Katayama, Y., Rudikoff, S., Pumphrey, J., Bowers, B., and Gottesman, S. (1990) *J. Biol. Chem.* **265**, 12536–12545
40. Liberek, K., Georgopoulos, C., and Zylicz, M. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 6632–6636
41. Zylicz, M., Ang, D., Liberek, K., and Zylicz, M. (1989) *EMBO J.* **8**, 1601–1608
42. Zylicz, M., Yamamoto, T., McKittrick, N., Sell, S., and Georgopoulos, C. (1985) *J. Biol. Chem.* **260**, 7591–7598
43. Tsurimoto, T., Hase, T., Matsubara, H., and Matsubara, K. (1982) *Mol. Gen. Genet.* **187**, 79–86
44. McMacken, R., Wold, M. S., LeBowitz, J. H. M Roberts, J. D., Mallory, J. B., Wilkinson, J. K., and Loehrlein, C. (1983) in *Mechanisms of DNA Replication and Recombination* (Cozzarelli, N., ed) pp. 819–848, Alan R. Liss, Inc., New York
45. Roberts, J. D., and McMacken, R. (1983) *Nucleic Acids Res.* **11**, 7435–7452
46. Rechsteiner, M. Hoffman, L., and Dubiel, W. (1993) *J. Biol. Chem.* **268**, 6065–6068
47. Squires, C., and Squires, L. C. (1992) *J. Bacteriol.* **174**, 1081–1085
48. Hwang, B. J., Park, W. J., Chung, C. H., and Goldberg, A. L. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 5550–5554
49. Lipinska, B., Zylicz, M., and Georgopoulos, C. (1990) *J. Bacteriol.* **172**, 1791–1797
50. Gottesman, S., Clark, W. P., de Crecy-Lagard, V., and Maurizi, M. R. (1993) *J. Biol. Chem.* **268**, 22618–22626

EXHIBIT 5

Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 8397–8401, September 1993
Biochemistry

FimC is a periplasmic PapD-like chaperone that directs assembly of type 1 pili in bacteria

C. HAL JONES*, JEROME S. PINKNER*, ANDY V. NICHOLES*, LYNN N. SLONIM*, SOMAN N. ABRAHAM*†, AND SCOTT J. HULTGREN*‡

*Department of Molecular Microbiology, Box 8230, Washington University School of Medicine, St. Louis, MO 63110; and †Department of Pathology and Medicine, Jewish Hospital of St. Louis, St. Louis, MO 63110

Communicated by Dale Kaiser, May 6, 1993

ABSTRACT **Biogenesis of the type 1 pilus fiber in *Escherichia coli* requires the product of the *fimC* locus. We have demonstrated that FimC is a member of the periplasmic chaperone family. The deduced primary sequence of FimC shows a high degree of homology to PapD and fits well with the derived consensus sequence for periplasmic chaperones, predicting that it has an immunoglobulin-like topology. The chaperone activity of FimC was demonstrated by purifying a complex that FimC forms with the FimH adhesin. A *fimC1* null allele could be complemented by the prototype member of the chaperone superfamily, PapD, resulting in the production of adhesive type 1 pili. The general mechanism of action of members of the chaperone superfamily was demonstrated by showing that the ability of PapD to assemble both P and type 1 pili was dependent on an invariant arginine residue (Arg-8), which forms part of a conserved subunit binding site in the cleft of PapD. We suggest that the conserved cleft is a subunit binding feature of all members of this protein family. These studies point out the general strategies used by Gram-negative bacteria to assemble adhesins into pilus fibers, allowing them to promote attachment to eukaryotic receptors.**

Periplasmic chaperones are part of a general secretory pathway required for the assembly of several well-characterized extracellular proteinaceous fibers, referred to as fimbriae or pili, in *Escherichia coli, Salmonella typhimurium, Salmonella enteriditis, Haemophilus influenzae, Bordetella pertussis, Yersinia pestis*, and *Klebsiella pneumoniae* (1–4). The best characterized and prototype member of the periplasmic chaperone superfamily is PapD from the P pilus system (1–9). Periplasmic chaperones stabilize pilus subunits in the periplasm through the formation of distinct periplasmic complexes (6, 8). The ability of chaperones to bind to and cap interactive surfaces on the pilus subunits prevents aggregation of the subunits and allows correct folding and assembly (6, 8).

The crystal structure of PapD has been solved to 2.0 Å resolution and revealed that PapD is composed of two domains each having an overall topology identical to an immunoglobulin fold (ref. 7; D. Ogg, personal communication). The two domains are connected by a hinge region and oriented such that a cleft is created between the two domains (7). Recent work by Slonim *et al.* (9) suggests that invariant surface-exposed residues that protrude into the cleft make up the subunit binding pocket of PapD. In contrast to cytoplasmic chaperones such as GroEL, DnaK, DnaJ, and SecB, which maintain their targets in highly unfolded conformations (10), PapD maintains target proteins in native-like conformations (8). In addition, periplasmic chaperones have an effector function, specifically targeting the subunits to outer membrane assembly sites for their incorporation into pili (11).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Type 1 pili are produced by nearly all Enterobacteriaceae (12). The major component of type 1 pili is repeating FimA subunits arranged in a right-handed helix to form a 7-nm-wide fiber with an axial hole (12). FimF, FimG, and FimH are three minor proteins associated with type 1 pili (13–15). FimH is the mannose-binding adhesin that promotes the interaction of type 1 piliated bacteria with mannose-containing glycoproteins on eukaryotic cell surfaces (14, 16). Type 1 pilus biogenesis requires two genes encoding nonstructural components of the pilus, which were originally identified as *pilB* and *pilC* (17). P pili, on the other hand, are produced specifically by pyelonephritic *E. coli* and mediate binding to Gal(α1-4)Gal-containing receptors on epithelial cell surfaces (1, 2). PapA is the major component of the P pilus rod while PapE, PapF, PapG, and PapK make up architecturally distinct fibers called tip fibrillae that are joined end-to-end to the pilus rods (18, 19). PapG is the Gal(α1-4)Gal binding adhesin, and presentation of this moiety in the fibrillum probably promotes efficient interaction with host receptors (1, 2). Assembly of the composite P pilus requires an immunoglobulin-like periplasmic chaperone and an outer-membrane usher, the products of the *papD* and *papC* genes, respectively (1, 2, 5, 11). As reported here, the FimC chaperone was found to have a similar structure, function, and mechanism of action as PapD. We suggest that all PapD-like chaperones utilize their immunoglobulin-like domains in a recognition paradigm to bind to pilus subunit proteins.

MATERIALS AND METHODS

Bacterial Strains. ORN103 (17) was used as a host in all complementation studies. HB101 (20) was used as a background strain for FimC purification. Induction of the *Pap* operon was as described (9). Induction of the type 1 operon was by three 48-hr static passages in Luria broth. Isopropyl β-D-thiogalactoside (IPTG) was used at 5 μM in liquid and 10 μM in solid media.

Hemagglutination Assay. Assays were performed following published protocols (9, 21). Soluble D-mannose (Sigma) at 0.1–0.5% was used to demonstrate mannose-sensitive hemagglutination (MSHA) while soluble Gal(α1-4)Gal at 0.1 mM was used to demonstrate specific hemagglutination by P pili.

Genetic Constructs. pPAP5 (5), pPAP37 (5), pPAP43 (22), pLS101 (9), pRT4 (23), and pJP1 (11) have been described. pJP3 is a pUC18 construct containing the complete type 1 operon cloned as a *Sal* I fragment from pSH2 (19). pJP5 contains an *Xho* I linker in the unique *Eco*RI site in pJP3, which defines the *fimC1* mutation. pJP4 is an *Eco*RI/*Hin*dIII subclone from pSJH9 (21) into the *tac* promoter plasmid pMMB91 (11). Sequencing of the *fimC* open reading frame§

Abbreviations: IPTG, isopropyl β-D-thiogalactoside; MSHA, mannose-sensitive hemagglutination.

‡To whom reprint requests should be addressed.

§The *fimC* sequence reported in this paper has been deposited in the GenBank data base (accession no. L14598).

with Sequenase (United States Biochemical) followed manufacturer's specifications.

Purification of FimC and Type 1 Pili. FimC was purified from the periplasm of HB101 containing pJP4. Bacteria were grown in a 5-liter fermentor (New Brunswick Scientific, Edison, NJ) to an A_{600} = 1.0 at which time IPTG was added to a final concentration of 1 mM. Induction was allowed to proceed for 1 hr, and then periplasm was prepared as described (9). Further purification of FimC by FPLC and HPLC was by a published protocol (5). Purified FimC protein was sent to Cocalico Biologicals (Reamstown, PA) for antibody production in rabbits. Electrophoresis of protein preparations was according to manufacturer's specifications (Bio-Rad). Western blot analyses followed published procedures (9). Purification and quantification of pili from equivalent gram quantities of cultures were as described (9).

Isolation of a FimC–FimH Complex. After induction with IPTG for 1 hr, periplasm from ORN103/pJP4 (*fimC*)/pRT4 (*fimH*) was prepared as described (9), dialyzed against phosphate-buffered saline and concentrated on a Centriprep column (Amicon) to 1.5 ml. Nine hundred microliters of the concentrated periplasm was mixed with 100 μl of washed mannose-Sepharose beads, rocked overnight, and then washed extensively with phosphate-buffered saline to remove unbound material. The FimC–FimH complex was eluted with 200 μl of 10% D-mannose or 10% methyl α-D-mannopyranoside (Sigma).

RESULTS

FimC Is an Immunoglobulin-Like Pilus Chaperone. Orndorff and Falkow (17) demonstrated that the product of the *fimC* gene is required for the assembly of type 1 pili by showing that a mutation in the *fimC* locus resulted in a nonpiliated phenotype (17). In addition, Hultgren *et al.* (21) found that the product of the *fimC* locus was required for the assembly of the mannose-binding adhesin on the surface of bacteria in the absence of the pilus rod. We have cloned (Table 1) and determined the complete nucleotide sequence (data not shown) of both strands of the *fimC* locus from the *E. coli* strain 149 that is a voided urinary isolate from a woman with cystitis (21, 24). The structural relationship between FimC and the superfamily of periplasmic chaperones was investigated by comparison of the predicted amino acid sequence of FimC and the chaperone consensus sequence (Fig. 1) (3, 4). This alignment showed that FimC was 32% identical and 50% homologous, considering conservative substitutions, to the prototype member of the immunoglobulin-like pilus chaperone family, PapD (Figs. 1 and 2). FimC contains all of the conserved residues that make up the hydrophobic core of pilus chaperones (Fig. 1); these residues participate in maintaining the overall structure of the two domains of the chaperone protein (3). FimC was also shown to possess the invariant internal salt bridge that has been proposed to be important in orienting the two domains of the chaperone with respect to one another (3). FimC also contains two invariant cleft residues (Arg-8 and Lys-112) that were shown by Slonim *et al.* (9) and M. Kuehn, D. Ogg, J. Kilberg, L.N.S., T. Bergfors, K. Flemmer, and S.J.H. (unpublished results) to make up a critical part of the subunit binding site of PapD (Figs. 1 and 2).

Table 1. Bacterial strains and plasmids used

Strain or plasmid	Genotype/remarks	Reference
Strain		
ORN103	*thr-1 leu-6 thi-1* Δ(*argF-lac*) *U169 xyl-7 ara-13 mtl-2 gal-6 rpsL tonA2 minA minB recA13* Δ*fim*(*EACDFGH*)	19
HB101	F⁻ *hsdS20* (r^-_B,m^+_K) *recA13 ara-14 proA2 lacY1 galK2 rpsL20* (Sm^r) *xyl-5 mtl-1 supE44*	22
Plasmid		
pPAP5	*Pap* operon, wild type	5
pPAP37	*papD1*, *Xho* I linker insertion	5
pLS101	*papD*, IPTG inducible	9
pJP1	*papD*, *papG*, IPTG inducible	12
pPAP43	P_{pap}, *papIBA*	24
pJP3	Type 1 operon, wild type	This study
pJP5	*fimC1*, *Xho* I linker insertion	This study
pJP4	*fimC*, IPTG inducible	This study
pRT4	*fimH*, IPTG inducible	25

The biochemical properties of FimC were also found to be characteristic of periplasmic pilus chaperones. Induction of *fimC* expression, cloned downstream of the *tac* promoter on pJP4 (Table 1), with IPTG resulted in the presence of a 26-kDa protein in the periplasmic space (Fig. 3*A*, lanes 1 and 2). When periplasmic extracts containing the induced 26-kDa protein were passed over a Mono S FPLC column, the FimC protein was eluted (95% homogeneity) in 0.15 M KCl (Fig. 3*A*, lane 3). FimC was purified to homogeneity by chromatography on a hydrophobic interaction column, CAA-HIC (Beckman), with conditions similar to that used to purify PapD for crystallography (data not shown). The amino-terminal sequence of the mature form of FimC was obtained, verifying that the induced 26-kDa band is the product of the FimC locus. FimC is a highly basic protein having an isoelectric point of 9.4 (data not shown). All of these properties are characteristic of periplasmic pilus chaperones (3, 8, 9).

Isolation of a FimC–FimH Complex. The chaperone activity of FimC was shown by demonstrating that FimC binds FimH, the type 1 adhesin, to form a periplasmic preassembly complex. Since FimH directs binding to receptors that contain mannose derivatives (14, 16, 23), mannose-Sepharose chromatography was used to purify a FimC–FimH complex

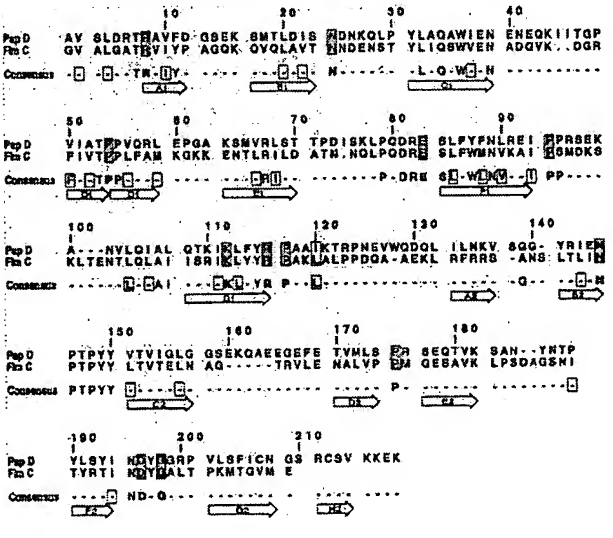


FIG. 1. The primary sequence of PapD and the inferred primary sequence of FimC are shown along with the chaperone consensus sequence, which was derived from 13 members of the family (3, 4). Open boxes in the consensus sequence indicate positions of conserved hydrophobic character. Boldfaced letters in the consensus sequence and stippled boxes in the PapD and FimC sequences indicate invariant residues while residues that are conserved in 8 out of 13 sequences are indicated in the consensus in normal type. Open arrows below the consensus sequence indicate the β strands identified on the crystal structure of PapD (7). This alignment is based on previously published data (4).

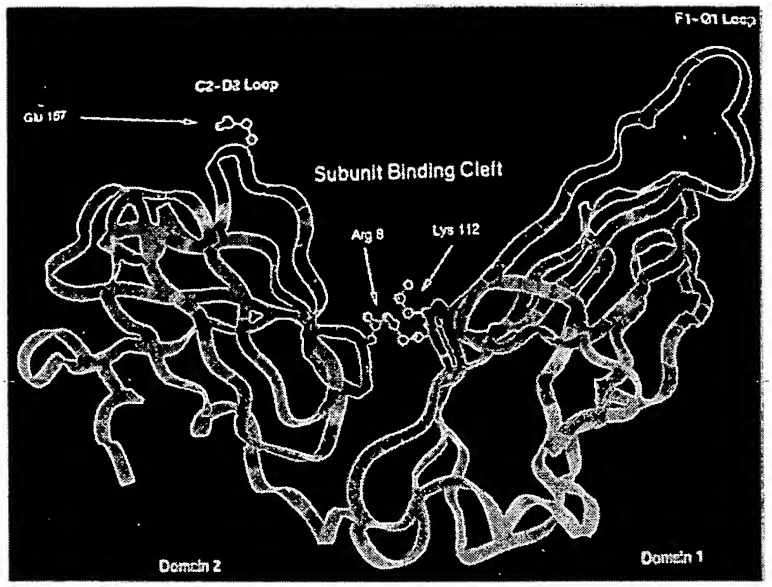


FIG. 2. Ribbon model of the crystal structure of PapD colored to indicate the homology of this protein with FimC. Identical residues are shown in purple while conservatively substituted residues are shown in blue. The side chains of the invariant residues Arg-8 and Lys-112 and the variable residue Glu-167 are indicated.

from the periplasm. Periplasm containing FimC and FimH (Fig. 3*B*, lane 1) was applied to mannose-Sepharose beads. A FimC–FimH complex was eluted with D-mannose (data not shown) and with methyl α-D-mannopyranoside (Fig. 3*B*, lane 3). Lane 4 in Fig. 3*B* shows that FimC alone is not retained by the mannose-Sepharose beads, but requires an interaction with FimH to be retained on the beads. We verified the identity of the eluted bands as FimC and FimH by Western blotting (data not shown) and by amino-terminal sequencing (Fig. 3 legend). Our results with the FimC–FimH complex demonstrate that FimH is folded in the complex such that the mannose-binding domain is in a native state and accessible for substrate binding. These results parallel those reported with the PapD–PapG complex (6, 8).

Structure–Function Relationship Between PapD and FimC. Given the predicted structural and observed biochemical relatedness between the chaperone proteins, we decided to test whether PapD could function in place of FimC in the assembly of type 1 pili. A knockout mutation in *fimC* (*fimC1*) on pJP5 abolished the ability of ORN103/pJP5 to produce pili as analyzed by EM (data not shown) and to mediate MSHA of guinea pig erythrocytes (Table 2). Remarkably, pLS101 (*papD*) complemented the *fimC1* null mutation (Table 2). Furthermore, we have demonstrated that both ORN103/pJP5 plus pLS101 (*papD*) and ORN103/pJP5 plus pJP4 (*fimC*) produce equivalent amounts of pili (data not shown). ORN103/pJP5 plus pLS101 (*papD*), however, had an 8-fold lower MSHA titer than ORN103/pJP5 plus pJP4 (*fimC*) (Table 2).

The ability of PapD to complement the *fimC1* genetic lesion confirms that the structures of PapD and FimC are highly related. The *fimC1* lesion, however, could not be complemented when PapG was coexpressed with PapD on plasmid pJP1 (*papD, papG*). ORN103/pJP5 (*fimC1*) plus pJP1 was MSHA-negative (Table 2) and nonpiliated as determined by EM (data not shown). PapD has been shown to bind to the PapG adhesin with a high affinity (8), which apparently is greater than the affinity that PapD has for the type 1 pilus

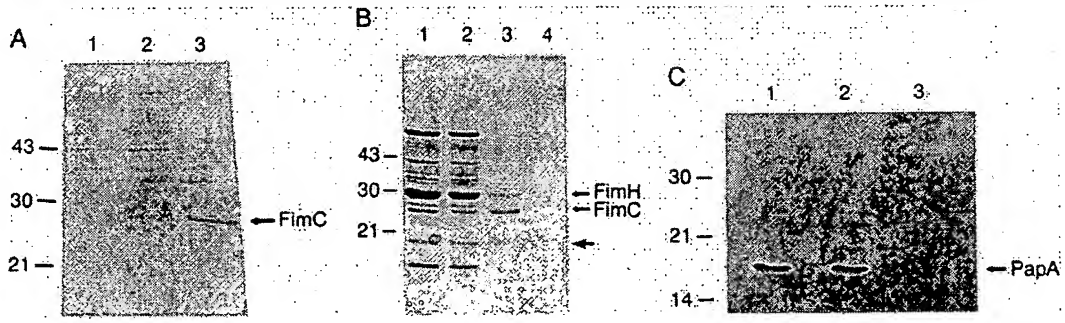


FIG. 3. Purification of FimC from the periplasm and subunit stabilization by the FimC chaperone. (*A*) HB101/pJP4 periplasm prepared from cells grown without (lane 1) and with (lane 2) 1 mM IPTG. A 26-kDa band, the predicted molecular mass for mature FimC, is seen only after addition of inducer. Lane 3 contains the eluate following chromatographic separation of induced periplasm as described (5). (*B*) Purification of the FimC–FimH complex. Lane 1 shows IPTG-induced periplasm from ORN103/pJP4 (*fimC*) plus pRT4 (*fimH*). The unbound material and the methyl α-D-mannopyranoside eluate from the mannose-Sepharose beads are shown in lanes 2 and 3, respectively. Shown in lane 4 is the eluate from induced ORN103/pJP4 (*fimC*) periplasm treated identically to the material in lane 3. The amino-terminal sequence of the bands in lane 3 was obtained by verifying the identity of the eluted proteins as FimH (FA–KTAN) and FimC (—AL–ATRVIY). The lower band in lane 3 (arrow) was identified by amino-terminal sequencing as a carboxyl-terminal truncate of FimH. (*C*) Western blot analysis with anti-PapA antiserum of periplasmic preparations showing the ability of FimC to stabilize the PapA subunit in the periplasm. In the absence of a chaperone, the PapA subunit is degraded in the periplasm (lane 3), whereas both PapD (lane 1) and FimC (lane 2) stabilize the subunit.

Table 2. Pilus assembly directed by a heterologous chaperone

Gene(s) induced in trans	Chaperone complementation		Test of dominance	
	Type 1* (*fimC1*)	Pap† (*papD1*)	Type 1* (*fimC*+)	Pap† (*papD*+)
None	3	0	850	256
fimC	850‡	0	512	190
papD	100‡	128	512	256
papD, papG	4	ND	768	ND
papD-R8A	0	0	ND	ND
papD-R8G	2	0	ND	ND
papD-R8M	0	0	ND	ND
papD-E167G	8	64	ND	ND
papD-E167T	8	96	ND	ND

Hemagglutination titer (9, 21) was used to measure pilus assembly and represents the dilution of bacteria that causes 50% hemagglutination of guinea pig erythrocytes for type 1 pili or human erythrocytes for P pili. Titers presented represent an average of at least three experiments. ND, not determined.

*Hemagglutination sensitive to 0.5% D-mannose.

†Hemagglutination sensitive to 0.1 mM Gal(α1-4)Gal.

‡Total amounts of type 1 pili and the relative ratio of FimH per milligram of FimA were measured and found to be equivalent in *papD*- and *fimC*-complemented strains.

proteins. As a consequence, coexpression of PapG with PapD apparently titrates PapD away from the type 1 pilus proteins, resulting in their misassembly. An alternative explanation for the MSHA-negative phenotype is that the PapD–PapG complex is targeted to the type 1 outer-membrane usher, FimD, but is not efficiently uncapped, resulting in a block in pilus assembly. We favor the former explanation since ORN103/pJP3 (wild-type type 1) plus pJP1 (*papD, papG*) has a wild-type MSHA-positive titer (Table 2), indicating that there is no dominant block to pilus assembly when PapD and PapG are coexpressed in a wild-type background.

Subunit Recognition in Heterologous Systems Requires the Chaperone Cleft. Slonim *et al.* (9) have shown that the highly conserved cleft region of PapD contains the subunit binding site with the invariant residue Arg-8 having a critical role in subunit recognition. We investigated the effect of three mutations in Arg-8 on the ability of PapD to modulate type 1 pilus assembly. Arg-8 was changed to alanine or glycine to completely remove the putative interactive Arg-8 side chain or to a methionine to maintain the bulk of the arginine side chain while removing the charge. None of these amino acid substitutions significantly affected the stability of PapD in the periplasmic space (9). All three mutations completely abolished the ability of PapD to assemble type 1 pili as well as P pili (Table 2), arguing that the recognition of both P and type 1 subunits by PapD requires the invariant Arg-8 cleft residue.

It has been suggested that the variable residues located in the loop structures that border the binding site cleft may play a role in chaperone specificity. One such residue in PapD is Glu-167, which is part of a negatively charged cluster of amino acids located in the loop that joins β strand C2 to D2 at the tip of domain 2 (3, 7, 9). Replacement of Glu-167 with glycine or threonine had only a modest effect on the ability of PapD to mediate P pilus assembly (Table 2) but significantly reduced the ability of PapD to direct type 1 pilus assembly, arguing for a role of this negatively charged region of the loop in interacting with type 1 subunits.

FimC Binds PapA But Does not Assemble P Pili. FimC was unable to complement the *papD1* lesion on pPAP37. ORN103/pPAP37 plus pJP4 (*fimC*) was hemagglutination-negative (Table 2) and nonpiliated as determined by EM (data not shown). Interestingly, the inability of FimC to mediate P pilus assembly was not due to the lack of interaction with the major subunit, PapA. A general property of P pilus subunits is that they follow nonproductive pathways that lead to aggregation and proteolytic degradation in the absence of an interaction with a periplasmic chaperone. Thus, PapA was proteolytically degraded when expressed from plasmid pPAP43 (23) in strain ORN103 (Fig. 3*C*, lane 3). Surprisingly, when PapA was coexpressed with either PapD or FimC, both chaperones were able to bind to PapA and stabilize the subunit in the periplasm so that it was detectable by Western blotting (Fig. 3*C*, lanes 1 and 2). The inability of FimC to direct the assembly of the Pap subunits into pili is presumably a reflection of the inability of FimC to orchestrate other critical interactions.

DISCUSSION

Periplasmic chaperones are a large family of proteins that are essential in the postsecretional assembly of higher order structures (1–9). Virtually nothing is known about periplasmic trafficking and targeting of proteins to the outer membrane; however, there is evidence that import of type 1 pilin subunits into the periplasmic space is dependent on the *sec* system (25). We suggest that in Gram-negative bacteria periplasmic chaperones are a component of a secretory pathway that also includes an outer-membrane transport protein (usher) to which the chaperone targets cognate proteins (11). Due to the interactive nature of pilus subunit proteins, this periplasmic transport system is probably required to receive nascently translocated subunits imported by the *sec* machinery. One defined role of periplasmic chaperones in this secretory pathway is to direct nascently translocated proteins down productive biological pathways by capping interactive surfaces to prevent aggregate formation and proteolytic degradation (8).

Previous work on the PapD chaperone, required for the assembly of P pili, has implicated the conserved cleft of the molecule in subunit binding (9). The recently determined crystal structure of PapD complexed with a PapG carboxyl-terminal peptide supported this conclusion by revealing that the peptide was anchored in the cleft by the invariant Arg-8 and Lys-112 residues and that it interacted along the G1 strand of PapD and with the F1 to G1 loop (M. Kuehn, D. Ogg, J. Kilberg, L.N.S., T. Bergfors, K. Flemmer, and S.J.H., unpublished results). Our demonstration of complementation of a *fimC1* null allele by PapD and the dependence of that activity on the Arg-8 residue suggests that the conserved cleft in periplasmic chaperones serves a common function in pilus biogenesis (Table 2). On the basis of the complementation results presented here, we suggest that the positively charged invariant cleft residues form a part of a universal anchor used by members of this protein family (Table 3) in chaperone–subunit interactions.

The differential effect of substitution at Glu-167 on assembly of P pili versus type 1 pili suggests a possible role for the C2–D2 loop in chaperone–subunit interactions (Table 2). The strong effect of the Glu-167 mutations on type 1 pilus assembly may reflect an involvement of the negatively charged region of the C2–D2 loop in interacting with the type 1 pilus subunits. This promotes the intriguing hypothesis that perhaps loops at the tips of both domains (Fig. 2) are important in chaperone–subunit interactions. In this model, a subunit would be anchored in the cleft by the invariant Arg-8 residue and make interactions along the faces of both domains of the chaperone.

Although *papD* complementation of the *fimC1* null allele produced type 1 pili and caused MSHA of guinea pig erythrocytes, the effective titer of these strains was 8-fold lower than the isogenic *fimC*-complemented strain (Table 2). Surprisingly, the lower hemagglutination titer was not due to decreased amounts of FimH-containing pili, since *fimC*- and

Table 3. Thirteen members of the chaperone superfamily

Chaperone	Mass, kDa	Organism	Surface structure	Homology to PapD*, %
PapD	28.5	*E. coli*	P pilus	
FimC	26	*E. coli*	Type 1 pilus	50
SfaE	23.5	*E. coli*	S pilus	62
FaeE	26.5	*E. coli*	K88 pilus	54
MrkB	25	*K. pneumoniae*	Type 3 pilus	60
HifB	25	*H. influenzae*	Pili	58
F17D	28	*E. coli*	F17 pilus	56
FanE	23	*E. coli*	K99 pilus	59
CS3-27kDa	27	*E. coli*	CS3 pilus	48
SefB	27	*S. enteriditis*	Fimbrin	55
FimB	24	*B. pertussis*	Types 2 and 3 pili	54
CaflM	26	*Y. pestis*	F1 envelope antigen	54
PsaB	25	*Y. pestis*	pH 6 antigen	57

The grouping of chaperones presented is based on published work (3, 4), except PsaB (from GenBank).

*Each protein was aligned to PapD, the prototype member of the chaperone superfamily, using the program GAP (version 7; Genetics Computer Group).

papD-complemented bacteria produced equivalent amounts of pili containing an equivalent amount of FimH per milligram of FimA (data not shown). In addition, using high-resolution EM, we discovered that type 1 pili assembled by either PapD or FimC had an identical composite architecture consisting of a short fibrillar tip structure joined end-to-end to the pilus rod (C.H.J. and S.J.H., unpublished results). Type 1 tip fibrillae are shorter than the previously described tip fibrillae of P pili (18, 19) but probably serve a similar function in presentation of the adhesin (2, 18, 19). Presumably, FimH was presented slightly differently in the PapD-assembled type 1 pili, which accounts for the lower hemagglutination titers, but the details of this "misincorporation" are not yet known.

We have shown that FimC is capable of binding to PapA in the periplasm (Fig. 3*C*) but fails to direct the assembly of adhesive P pili (Table 2). One explanation for these findings is that the FimC chaperone fails to interact with PapF and PapK. PapF and PapK are two specialized subunits required to initiate the assembly of the tip fibrillum and pilus rod and to join each structural element within the pilus (18, 19). A *papF, papK* double mutant abolishes the ability of the bacteria to produce pili (19). Although FimC binds to PapA, it may be unable to bind to other Pap subunit proteins; we have shown that FimC, although it interacts with PapG, fails to form a FimC–PapG complex that is competent to bind to receptor (data not shown). Alternatively, the FimC–subunit complexes may fail to interact with PapC, the outer membrane usher in the P pilus system (11). As with the entire chaperone family, PapD and FimC diverge considerably in the loop regions and as in domain 2 of the molecule (Figs. 1 and 2). These regions might define the specificity of the interaction of the chaperone–subunit complex with the usher.

Our studies on the function of two chaperones in pilus biogenesis allow for the formulation of a model that describes the interaction of the chaperone with subunit proteins. The subunits probably anchor to the invariant Arg-8 residue deep in the chaperone cleft and interact along both cleft faces of domain 1 and domain 2 including the variable loops at the tips of each domain. The ability to understand the fine molecular details of the mechanism of action of molecular chaperones in the development of pilus organelles will continue to unveil general principles of macromolecular assembly. Crystals of FimC that defract to 3.2 Å have already been generated from the purified FimC obtained from these studies, and the three-dimensional structure is currently being solved (S. Knight, personal communication).

The work described was supported by grants to S.J.H. from the Lucille P. Markey Charitable Trust Contract and National Institutes of Health Grant 1RO1AI29549. S.N.A. is supported by funds from Public Health Service Grant AI-13550 from the National Institutes of Health. C.H.J. is supported by a W. M. Keck fellowship.

1. Hultgren, S. J., Normark, S. & Abraham, S. N. (1991) *Annu. Rev. Microbiol.* **45**, 383–415.
2. Jones, C. H., Jacob-Dubuisson, F., Dodson, K., Kuehn, M., Slonim, L., Striker, R. & Hultgren, S. J. (1992) *Infect. Immunol.* **60**, 4445–4451.
3. Holmgren, A., Kuehn, M. J., Branden, C.-I. & Hultgren, S. J. (1992) *EMBO J.* **11**, 1617–1622.
4. Jacob-Dubuisson, F., Kuehn, M. J. & Hultgren, S. J. (1993) *Trends Microbiol.* **1**, 50–55.
5. Lindberg, F., Tennent, J. M., Hultgren, S. J., Lund, B. & Normark, S. (1989) *Proc. Natl. Acad. Sci. USA* **84**, 5989–5902.
6. Hultgren, S. J., Lindberg, F., Magnusson, G., Kilberg, J., Tennent, J. M. & Normark, S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4357–4361.
7. Holmgren, A. & Branden, C.-I. (1989) *Nature (London)* **342**, 248–251.
8. Kuehn, M. J., Normark, S. & Hultgren, S. J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10586–10590.
9. Slonim, L. N., Pinkner, J. S., Branden, C.-I. & Hultgren, S. J. (1992) *EMBO J.* **11**, 4747–4756.
10. Lecker, S. H., Lill, R., Ziegelhoffer, T., Georgopoulos, C., Bassford, P., Kumamoto, C. A. & Wickner, W. (1989) *EMBO J.* **8**, 2703–2709.
11. Dodson, K. D., Jacob-Dubuisson, F., Striker, R. T. & Hultgren, S. J. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3670–3674.
12. Brinton, C. C. (1965) *Trans. N. Y. Acad. Sci.* **27**, 1003–1065.
13. Maurer, L. & Orndorff, P. E. (1987) *J. Bacteriol.* **169**, 640–645.
14. Abraham, S. N., Goguen, J. D., Sun, D., Klemm, P. & Beachey, E. A. (1987) *J. Bacteriol.* **169**, 5530–5536.
15. Russel, P. W. & Orndorff, P. E. (1992) *J. Bacteriol.* **174**, 5923–5935.
16. Ofek, I., Mirelman, D. & Sharon, N. (1977) *Nature (London)* **265**, 623–625.
17. Orndorff, P. E. & Falkow, S. (1984) *J. Bacteriol.* **159**, 736–744.
18. Kuehn, M. J., Heuser, J., Normark, S. & Hultgren, S. J. (1992) *Nature (London)* **356**, 252–255.
19. Jacob-Dubuisson, F., Heuser, J., Dodson, K., Normark, S. & Hultgren, S. J. (1993) *EMBO J.* **12**, 837–847.
20. Boyer, H. W. & Roulland-Dussoix, D. J. (1969) *Mol. Biol.* **41**, 459–472.
21. Hultgren, S. J., Duncan, J. L., Schaeffer, A. J. & Amundsen, S. K. (1990) *Mol. Microbiol.* **4**, 1311–1318.
22. Lund, B., Lindberg, F. P., Baga, M. & Normark, S. (1987) *J. Bacteriol.* **162**, 1293–1301.
23. Tewari, R., MacGregor, J. I., Ikeda, T., Little, J. R., Hultgren, S. J. & Abraham, S. N. (1993) *J. Biol. Chem.* **268**, 3009–3015.
24. Hultgren, S. J., Porter, T. N., Schaeffer, A. J. & Duncan, J. L. (1985) *Infect. Immunol.* **50**, 370–377.
25. Dodd, D. C., Bassford, P. J., Jr., & Eisenstein, B. I. (1984) *J. Bacteriol.* **159**, 1077–1079.

EXHIBIT 6

Microbial Cell Factories

BioMed Central

Review

Open Access

Protein quality control in the bacterial periplasm

Marika Miot and Jean-Michel Betton*

Address: Unité Repliement et Modélisation des Protéines, Institut Pasteur, CNRS-URA2185, 28 rue du Dr Roux, 75754 Paris cedex 15, France

Email: Marika Miot - mmiot@pasteur.fr; Jean-Michel Betton* - jmbetton@pasteur.fr

* Corresponding author

Published: 07 May 2004

Microbial Cell Factories 2004, 3:4

Received: 18 March 2004

Accepted: 07 May 2004

This article is available from: http://www.microbialcellfactories.com/content/3/1/4

Abstract

The proper functioning of extracytoplasmic proteins requires their export to, and productive folding in, the correct cellular compartment. All proteins in *Escherichia coli* are initially synthesized in the cytoplasm, then follow a pathway that depends upon their ultimate cellular destination. Many proteins destined for the periplasm are synthesized as precursors carrying an N-terminal signal sequence that directs them to the general secretion machinery at the inner membrane. After translocation and signal sequence cleavage, the newly exported mature proteins are folded and assembled in the periplasm. Maintaining quality control over these processes depends on chaperones, folding catalysts, and proteases. This article summarizes the general principles which control protein folding in the bacterial periplasm by focusing on the periplasmic maltose-binding protein.

Review

Although all the information necessary for a protein to reach a native structure is contained in its amino acid sequence, *in vivo* protein folding requires the participation of protein factors; molecular chaperones, folding catalysts and proteases, that monitor or regulate this process through many cellular functions. Under physiological conditions, these factors maintain quality control of protein biosynthesis, and errors or failures of the protein folding process are rare. However, incorrectly folded or misfolded proteins can appear as a result of (i) spontaneous or inducible mutations that affect protein folding pathways, (ii) exposure of the cell to environmental stress, such as elevated temperatures or hyperosmolarity, and (iii) overexpression of recombinant genes. In these cases, the polypeptide chain, instead of folding into the native biologically active state, misfolds, and eventually induces a stress response [1]. The resulting misfolded proteins may either be degraded by proteases, repaired by chaperones, or aggregated and sequestered as inclusion bodies (IBs), when escaping protein folding quality control [2]. In the cytoplasm of *Escherichia coli*, this control is performed mainly by a set of stress-inducible chaperones and proteases collectively known as heat shock proteins (Hsps). The presence of cellular membrane-bounded compartments in *E. coli* further complicates these issues of protein biosynthesis, and implies the existence of distinct folding and degradation machineries, quality control systems, for extracytoplasmic proteins [3]. The bacterial periplasm is separated from the extracellular milieu only by the porous outer membrane, and hence is more susceptible to changes in the environment than the cytoplasm. Thus, protein quality control in the periplasm is crucial for the survival of bacterial cells. Here we summarize the main features of protein folding in the periplasm of *E. coli*, and focus on studies using a model system, the maltose-binding protein.

Biosynthesis of periplasmic proteins

In *E. coli*, exported proteins with an ultimate destination of the periplasm and outer membrane are synthesized as precursors with a cleavable amino-terminal signal

sequence. Depending on the nature of the precursors, different pathways exist for their transport across the inner membrane. Although the translocation of most exported proteins occurs by the general secretory (Sec) pathway [4], across a channel formed by the heterotrimeric membrane complex SecYEG, a subset of precursors, containing two consecutive arginine residues near the N-terminus of the signal sequences, utilizes the twin-arginine translocation (Tat) pathway. The Tat machinery, encoded by the *tat*ABCE genes, exports precursors that must completely fold in the cytoplasm to bind cofactors [5].

Targeting of precursors to the SecYEG translocase

In marked contrast, the Sec pathway transports proteins which are almost unstructured, as it was confirmed by the recently solved crystal structure of the SecY complex from *Methanococcus janansahii* [6]. Since SecYEG forms a passive channel, it must be associated with other components to provide the driving force for translocating preprotein substrates. According to their degree of coupling to translation, two alternative modes of protein translocation have been identified. In cotranslational translocation, as the signal sequence emerges from a translating ribosome, it binds to the signal recognition particle (SRP), and, in concert with its membrane-bound receptor FtsY, the nascent polypeptide is targeted to SecYEG. The energy for translocation comes from GTP hydrolysis during translation. In posttranslational translocation, many newly synthesized precursors interact with the chaperone SecB (see below). Then, both SecB and preproteins provide binding sites for SecA, a peripheral membrane bound ATPase, that targets these initial complexes to the translocase. By mediating repeated cycles of ATP binding and hydrolysis, SecA pushes the polypeptide chain through the SecYEG channel [7]. The combined data available suggest that the SRP-dependent pathway is probably mainly involved in the cotranslational assembly of cytoplasmic membrane proteins [8,9], and that most periplasmic and outer membrane proteins are targeted posttranslationally to the SecA/SecB pathway [9]. However, recent studies indicate that some extracytoplasmic proteins exhibit a combined dependence on SRP/FtsY and SecA [10,11]. Even though protein translocation in *E. coli* has been well studied, the molecular mechanisms which allow the discrimination between both pathways converging at the SecYEG translocase remain unknown.

When a nascent chain emerges from the ribosome, a range of folding and targeting factors are waiting to bind the precursor. The first folding helper encountered is probably the trigger factor [12], a ribosome-bound peptidyl-prolyl isomerase [13] which displays chaperone activity *in vitro*. In addition to trigger factor, the emerging signal sequence is accessible to SRP and SecA [14]. These initial contacts are then modified upon further growth of the nascent chain and binding to SecYEG. In the Sec pathway, during elongation or after release from the ribosome, many, but not all, polypeptide chains bind to SecB, a tetrameric chaperone which stabilizes the precursor in a translocation-competent conformation [15]. The SecB-bound precursors, in cooperation with SecA, are then targeted to SecYEG. Although other cytoplasmic chaperones, like DnaK (Hsp70) and GroEL (Hsp60), may participate in protein export by preventing early folding of precursors [16,17], these general chaperones have a marginal role in targeting precursors to the Sec translocase. However, when translocation is impaired, by slowly exported proteins or defective signal peptides for example, precursors could accumulate in the cytoplasm. Structurally, a precursor in an export-competent conformation may be similar to an unfolded protein. Indeed, Wild et al. [17] showed that accumulation of precursors in strains lacking SecB induces a stress response mediated by the specific sigma factor, σ^{32}. Under normal conditions, DnaK and DnaJ sequester σ^{32} through direct binding, inhibiting its association with RNA polymerase [18]. Accumulation of unfolded precursors displaces this complex and allows σ^{32} binding to RNA polymerase to direct transcription of heat shock genes. It has been suggested that the final distribution of the precursors among the different pathways is determined by a kinetic partitioning that is dependent on the rate of folding, relative to the rate of chaperone binding [19]. Finally, how these different factors compete will determine whether initiation of translocation occurs efficiently (Figure 1).

In the next step, the precursor inserts as a loop into the SecYEG channel, with its signal sequence intercalated between two transmembrane helices of SecY, and with its mature region in the pore. The polypeptide chain is then transported through the pore, and the signal sequence is cleaved at some point during translocation by the signal peptidase, a membrane-bound endopeptidase with its catalytic domain facing the periplasm [20].

Protein folding in the periplasm

Once they reach the periplasm and after signal sequence cleavage, periplasmic proteins may encounter two types of protein folding catalysts: protein disulphide isomerases (Dsb proteins), which catalyse disulfide-bond formation, and peptidyl-prolyl isomerases (PPIase), which catalyse *cis-trans* isomerization of peptidyl bonds. These enzymes are known to catalyze rate-limiting steps in the *in vitro* protein folding process.

The formation of disulfide bonds is essential for the oxidative folding, activity, and stability of many proteins exported from the cytoplasm. Disulfide bonds formation results from elaborate electron transfer pathways between the Dsb oxydoreductases [21]. DsbA is the strongest thiol

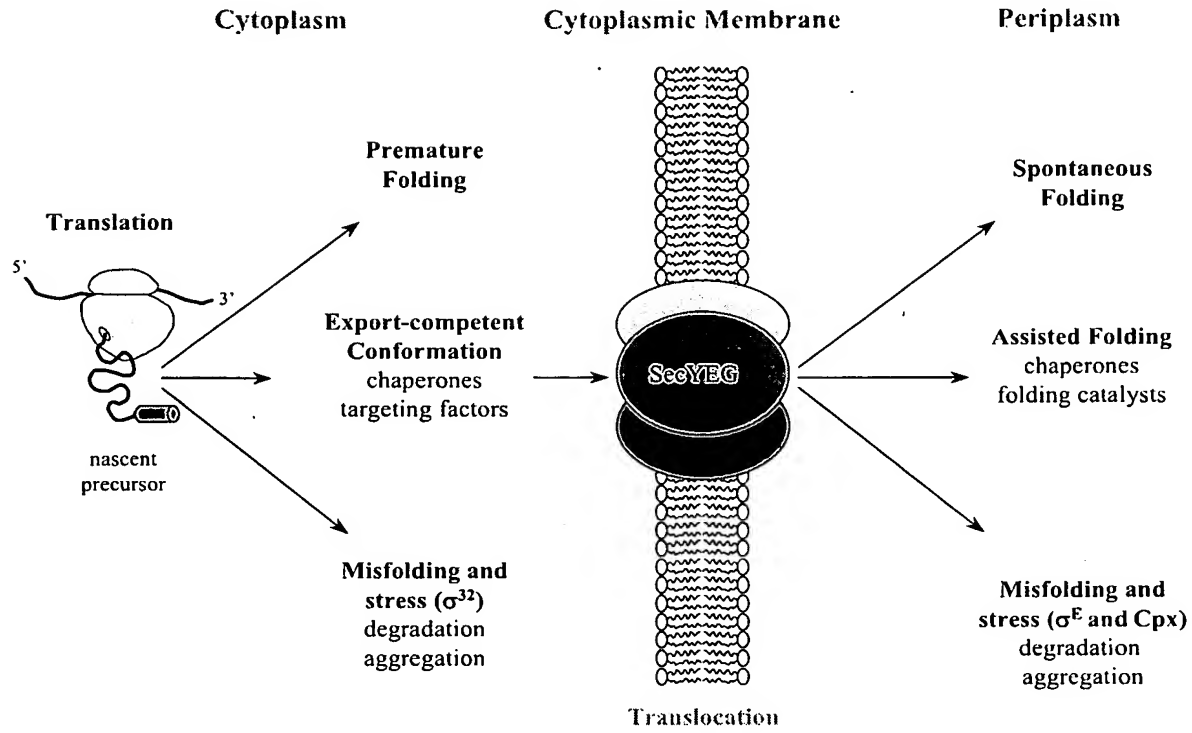


Figure 1
Alternative folding pathways for periplasmic protein in E. coli. This scheme illustrates the present discussion and tries to emphasize the alternative and competing process between folding and misfolding of periplasmic proteins before and after translocation across the cytoplasmic membrane.

oxidant which catalyses oxidation of cysteines of folding proteins. This reaction occurs via formation of mixed disulfide complexes between DsbA and its substrates. DsbB, a membrane-bound enzyme, maintains DsbA oxidized by transferring electrons from reduced DsbA to quinones in the respiratory chain [22]. It has become clear that disulfide bond formation, far from being perfect, occurs between incorrect pairs of cysteines resulting in the presence of non-native disulfides in proteins. As periplasmic thiol isomerases, DsbC and DsbG promote the rearrangements of incorrect disulfide bonds. The maintenance of these thiol isomerases in the reduced state depends on the inner membrane protein DsbD, which utilizes the thioredoxin/thioredoxin reductase system in the cytoplasm as a source of reducing equivalents [23]. Recently, covalent intermediates between DsbA and exported substrate proteins were detected *in vivo* by using a DsbA variant carrying the $Pro^{151} \rightarrow Thr$ substitution in its active site [24]. This study opens the way to understand at what point during protein translocation and folding, disulfide bond formation takes place in the periplasm.

Representatives of the three unrelated families of PPIases have been identified in the periplasm: PpiA (also known as CypA or RotA) is a cyclophilin that is not inhibited by cyclosporin A [25]; FkpA is related to the FK506-binding protein family, or FKBP, [26]; and PpiD [27] and SurA [28], two parvulin homologs. Although all PPIases catalyze the rate-limiting step in the refolding of RNase T1 [29]*in vitro*, their cellular role remains enigmatic. With the exception of *surA*, the absence of a significant phenotype for null mutants indicates that these periplasmic PPIases are not essential for viability or have overlapping functions.

Aside from specific chaperones, such as PapD involved in pilus assembly [30], few periplasmic chaperones have been identified, and there are no classical Hsp chaperones, such as DnaK or GroEL, in this compartment. Indeed, as the periplasm lacks ATP, periplasmic chaperones must be mechanistically distinct from their cytoplasmic counterparts, most of which use ATP to drive their cycles of substrate binding and release. However, global searches using genetic selections based on σ^E activity (see below) have identified the *surA* and *skp/ompH* genes [28,31], which are involved in the folding and assembly of outer membrane proteins [32]. Furthermore, *in vitro* chaperone activity of several periplasmic proteins such as DsbC [33] and DsbG [34] or substrate-binding proteins [35,36] have already been reported, but their contribution has not yet been studied *in vivo*. It should be noted that studying the *in vivo* function of a putative chaperone is more difficult than assessing its activity *in vitro*, using a classical chaperone substrate. However, it is tempting to speculate from these observations that, because of their unique active sites, protein folding catalysts can perform chaperone functions in order to compensate for the apparent deficiency of general chaperones in the periplasm. Indeed, FkpA and SurA exhibit PPIase-independent chaperone activity *in vivo* [37,38]. Although these two PPIases are unrelated, their crystal structures share a similar organization: a chaperone activity residing in the N-terminal domains and the C-terminal catalytic domains tethered ~30Å apart, connected either by a long α helix [39], or an extended polypeptide [40]. In other foldases that exhibit both chaperone and protein folding catalytic activities, these two functions have also been assigned to distinct regions or domains of the molecule. For instance, the trigger factor comprises three domains, with the central FKBP-like domain displaying PPIase activity and the C-terminal domain being implicated in chaperone activity [41]. The structures of DsbC and FkpA present also further striking similarities. Both enzymes are homodimeric, with subunits divided into two domains where the N-domains form the dimeric interfaces. The C-domains, carrying the catalytic sites, are connected to the N-domains by a pliable helix, conferring a V-shaped form on the dimers, and orienting the two catalytic sites face inwards towards each other. The cleft formed in the DsbC dimer by association of the N-domains is hydrophobic in character, and it was proposed that this region forms the binding site of the unfolded polypeptide for the chaperone activity, sequestering the substrate from aggregation and, at the same time, giving it access to the catalytic sites to facilitate disulfide bond exchange [42]. In FkpA, the cleft formed by association of the N-domains has approximately the same dimensions to that formed in DsbC. In contrast to the hydrophobic nature of the DsbC cleft, that of FkpA has an intermediate negative potential, reflecting probably a difference in their substrate specificity. However, it is an attractive idea that the V-shaped form of these foldases cradles the unfolded polypeptide substrates while giving access to the isomerase catalytic sites. The innate flexibility of the C-domains would allow for adaptability between the chaperone and isomerase activities.

An alternative fate for misfolded periplasmic proteins is degradation by proteases (Figure 1). Although more than 10 periplasmic proteases have been identified in *E. coli*, DegP (also known as HtrA or Do) is the only protease identified as a heat shock protein involved in the degradation of misfolded proteins [43,44]. Indeed, *degP*, which is essential for the survival of bacteria above 40°C, is under the transcriptional regulation of the alternative heat shock sigma factor, σ^E [45]. The reason for this essential role is not clear, but presumably an increased number of misfolded proteins accumulate under these conditions. DegP is a widely conserved serine protease composed of an N-terminal domain believed to have regulatory functions, a trypsin-like protease domain, and two PDZ domains [46]. PDZ domains are protein modules that mediate specific protein-protein interactions. The crystal structure of DegP reveals a staggered association of two trimeric rings forming a hexamer [47]. The proteolytic sites are located in a central cavity that is accessible only laterally from the PDZ domains, which probably recognize and bind substrates. It has been shown that the activity of DegP could switch between chaperone and protease activities in a temperature-dependent manner [48]. At temperatures below 28°C, DegP acts as a chaperone, protecting misfolded proteins from irreversible aggregation, and above 28°C, its protease activity increases dramatically degrading misfolded proteins. However, *in vivo* DegP which does not act as a general chaperone in the periplasm [49], is the major proteolytic machine responsible for periplasmic protein quality control.

Extracytoplasmic stress response

Because optimal cellular growth requires that the cell be able to sense and respond to changes in subcellular compartments, it is not surprising that the stress response in Gram-negative bacteria is compartmentalized into cytoplasmic and extracytoplasmic responses. A common feature of inducers of the *E. coli* extracytoplasmic stress response is their potential to generate excessive amounts of misfolded proteins in the envelope, particularly outer membrane proteins, and many of the proteins involved in periplasmic protein quality control have been characterized from studies on the extracytoplasmic stress responses.

In contrast to cytoplasmic stress, where the sensing of misfolded proteins and the accompanying response take place in the same compartment, extracytoplasmic stress signals must cross the cytoplasmic membrane by a signal

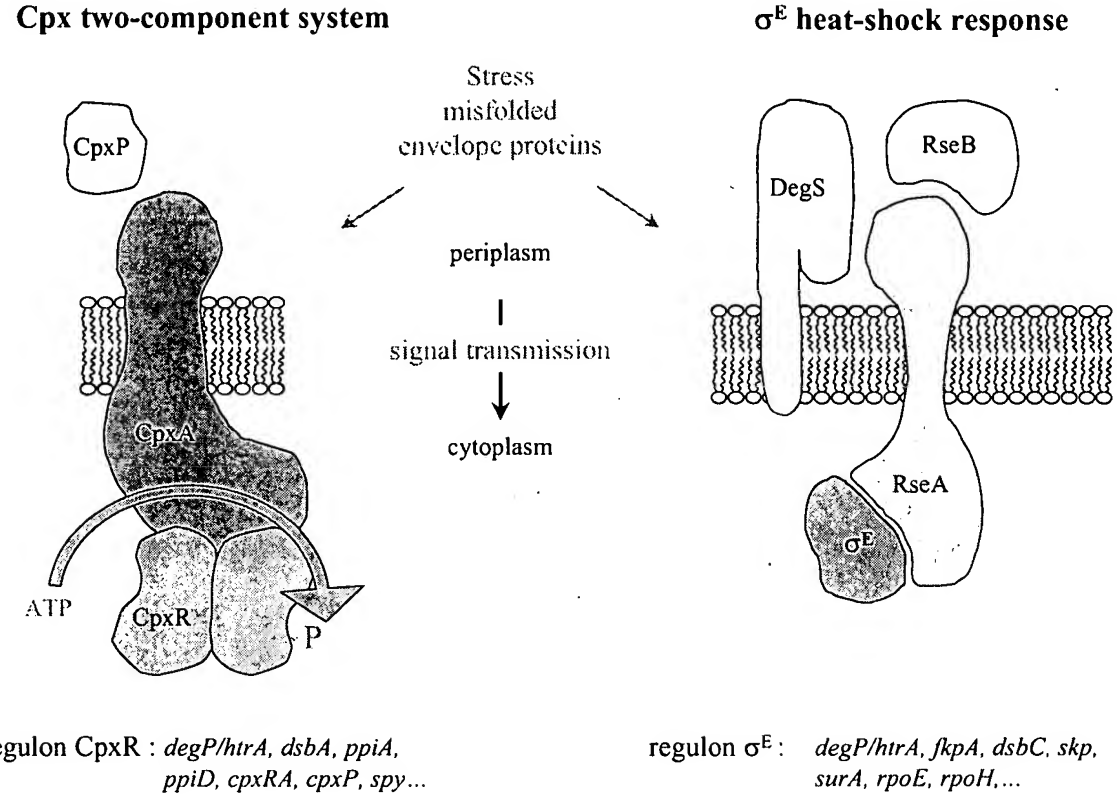


Figure 2
Two signaling pathways for extracytoplasmic stress responses in E. coli. Components of these pathways reside on each side of the cytoplasmic membrane (periplasm and cytoplasm). In the presence of misfolded periplasmic proteins, an unknown mechanism triggers a conformational change of the CpxA that leads to an activation of its kinase activity and promote an increase in CpxR phosphorylation. Under normal physiological conditions, σ^E is sequestred by RseA and DegS is inactive, with its PDZ domain inhibiting the protease activity. Recognition by the PDZ domain of misfolded periplasmic proteins relieves inhibition of the DegS protease. Activated DegS cleaves RseA, which releases σ^E. For both signal transduction systems, a partial list of the downstream targets is indicated.

transduction system. *E. coli* senses and responds to extracytoplasmic stress via at least two overlapping, but distinct, transduction pathways, the Cpx two-component system and the heat shock σ^E pathway (Figure 2). Both regulatory systems control the expression of several genes whose products are envelope-localized protein folding catalysts, chaperones and proteases, as well as genes involved in lipid and lipopolysaccharide metabolism. Together, these proteins serve to ensure proper biogenesis of the bacterial envelope by preventing any perturbation in periplasmic protein folding [50].

The Cpx pathway is a typical two-component signal transduction system, consisting of the hisitidine kinase CpxA and the response regulator CpxR [51]. The input signal is transduced across the cytoplasmic membrane by the sensor kinase CpxA, which is composed of a periplasmic domain and a conserved cytoplasmic signaling domain separated by two transmembrane α-helices [52]. Upon detection of envelope stress, CpxA autophosphorylates by using ATP at a conserved histidine and then transfers this phosphate to a conserved aspartate in the N-terminal domain of CpxR. Phosphorylated CpxR activates transcription of genes whose products are involved in envelope physiology. The Cpx pathway is induced by elevated

pH [53], altered inner membrane composition [54], and overproduction of envelope proteins like the outer membrane lipoprotein NlpE [55] or pili subunits [56,57]. The Cpx regulon is positively autoregulated and is feedback inhibited by CpxP, a small periplasmic protein [58]. The response seems to be modulated by direct interaction of the periplasmic domain of CpxA, probably in concert with CpxP [59]. Some of the targets regulated by CpxR are the heat shock protease DegP [60], the peptidyl prolyl cis/trans isomerases PpiA [61] and PpiD [27], and the disulfide oxidoreductase DsbA [61]. Apart from the extracytoplasmic stress response the Cpx system is implicated in conjugation [62], invasion of host cells [63], and biofilm formation [64,65].

The σ^E envelope stress response is induced by heat shock [66,67] and by perturbations to outer membrane protein folding [31,68]. Under nonstress conditions, the activity of σ^E is negatively regulated by the inner membrane anti sigma factor RseA and by the periplasmic protein RseB [69-71]. In response to extracytoplasmic stress, RseA is rapidly degraded by the sequential action of DegS [72] and YaeL/EcfE proteases [73] that frees σ^E to associate with RNA polymerase and directs the transcription of its regulon. Recently, it was shown that peptides mimicking the C-terminal sequence of outer membrane proteins bind the DegS PDZ domain, activate DegS cleavage of RseA, and induce σ^E-dependent transcription [74]. These data suggest that DegS acts as a sensor of extracytoplasmic stress by binding unassembled outer membrane proteins, and that DegS activation involves relief of inhibitory interactions between its PDZ and protease domain. However, RseB, which interacts with misfolded periplasmic proteins [71], might influence the accessibility of RseA to cleavage by DegS. At least 20 genes were identified to be regulated by σ^E [75] which include *rpoH*, *degP* [45], *yaeL/ecfE* [76], and *fkpA* [77].

Focus on the maltose-binding protein

To study protein folding in the periplasm of *E. coli*, we use a model system, the maltose-binding protein (MalE or MBP). MalE is the soluble receptor for the high affinity transport of maltose and maltodextrins into the periplasm [78]. Because of its particular function in maltose transport, correct export and folding of MalE in this cellular compartment is essential for cells to utilize maltose as a carbon source [79]. This feature has facilitated the development of powerful genetic selections or of screens which led to the discovery of several genes composing the Sec pathway [80]. MalE has also been extensively used as a model of protein translocation [81]. Translocation of MalE requires that the nascent preMalE polypeptide chain reach a critical size corresponding to 80 % of the final length, and that the precursor exist in an export-competent conformation representing a partially unfolded state [82]. SecB binding mainly participates in the maintenance of this initial conformation. In addition to this chaperone activity, SecB is an important targeting factor for MalE. Indeed, when the signal sequence is deleted, translocation of MalE becomes entirely SecB dependent [83]. Then, MalE is translocated across the SecYEG channel, released into the periplasm after signal sequence cleavage, and the newly exported MalE folds into its native structure. The crystal structure of MalE consists of two discontinuous domains constructed from secondary structural βαβ units surrounding a cleft that forms the binding site for maltose [84]. These properties make MalE an ideal model with which a direct comparison of the mechanism of *in vitro* and *in vivo* folding is greatly facilitated.

MalE31, a model for periplasmic protein misfolding

For many years, and well before the resolution of its 3D-structure, the laboratory of M. Hofnung studied the MalE structure-function relationships. One genetic approach, using random insertion of a *BamH*I oligonucleotide linker, led to the identification of the so-called 'permissive sites' mutants of *malE* in which insertions or deletions were structurally tolerated without loss of biological properties [85]. In contrast, other random insertions in *malE* prevented both growth on maltose and periplasmic release of the corresponding proteins [86]. In many cases, a large deletion of several amino acid residues from the mature region of the protein could explain their defective phenotypes. However, in one case, *malE31*, the linker insertion did not modify the length of the protein and resulted in $Gly^{32} \rightarrow$ Asp and $Ile^{33} \rightarrow$ Pro substitutions. Structurally, these residues are located in a turn connecting the helix αI to strand βB of the first βαβ element within the N-domain (Figure 3A). The $Ile^{33} \rightarrow$ Pro substitution was identified as the major factor responsible for the Mal$^-$ phenotype [87]. Indeed, cells expressing a *malE* variant carrying the $Gly^{32} \rightarrow$ Asp substitution displayed biological properties similar to those of cells expressing wild-type *malE*. The subcellular location of these mutants revealed that MalE31 was entirely recovered in the membrane fraction. This result suggested that export of MalE31 was blocked somewhere in the membrane or, alternatively, that MalE31 formed inclusion bodies (IBs) either in the cytoplasm from the precursor, before translocation, or in the periplasm, from the mature protein, after release from the translocase. It is technically very difficult to rigorously differentiate between these cellular fates by cell-fractionation experiments only. Thus, to determine whether insoluble MalE31 which co-sedimented with membranes was consistent with aggregated proteins rather than membrane-associated proteins, a crude lysate was fractionated by flotation gradient centrifugation, and revealed that MalE31 remained near the bottom [88]. Although cell fractionation from spheroplasts is less sensitive to artifacts than the osmotic shock procedure, both

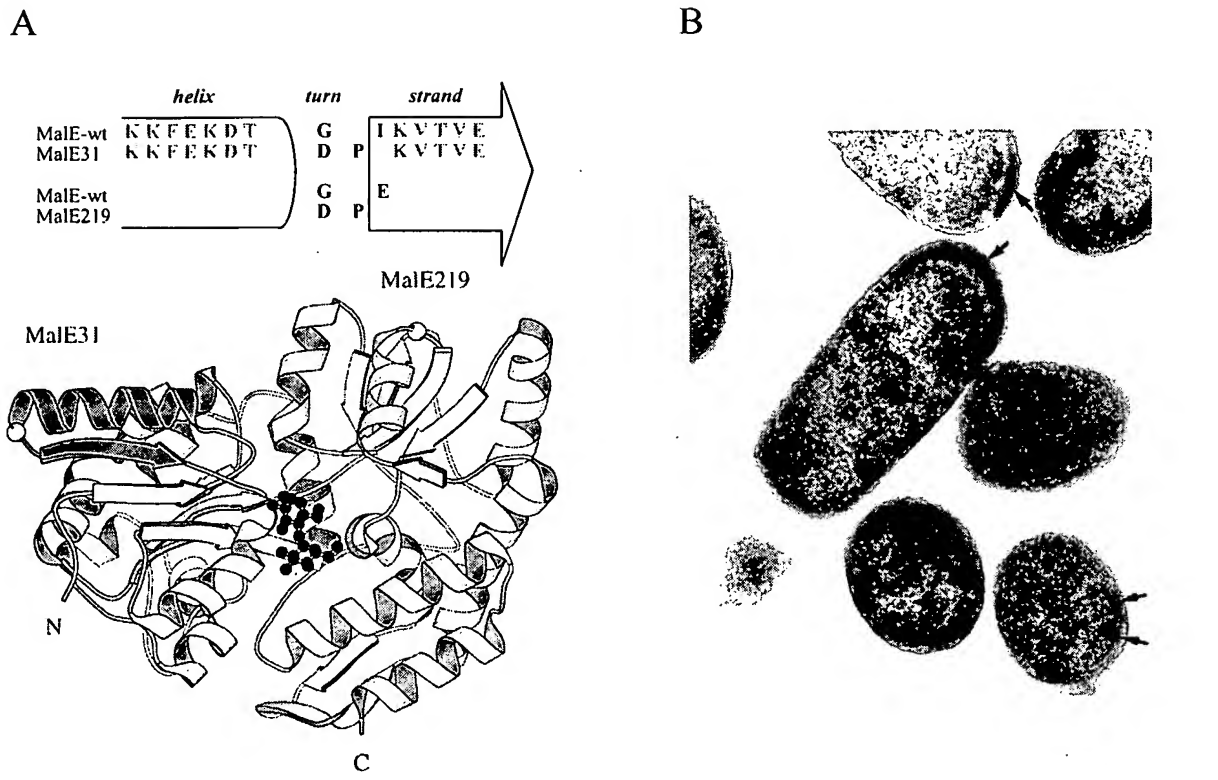


Figure 3
In vitro and in vivo fates for MalE31. A- The maltose-bound state of MalE31. The helix αI and strand βB of the N-domain (MalE31), colored in magenta, and the helix αVII and strand βJ of the C-domain (MalE219), colored in blue, are located in the crystal structure of MalE. For both turns, the Gly residue is shown in yellow. The bound maltose substrate in black is located in a cleft between both domains. B- Periplasmic inclusion bodies of MalE31. Transmission electron micrograph of ultrathin sections of *E. coli* overproducing *malE31* at 30°C. The black arrows indicate the position of periplasmic inclusion bodies.

methods cannot discriminate between cellular location (cytoplasmic or periplasmic) of IBs. However, processing of the precursors, a late step of export (see above), is a reliable indicator that the protein has, at least partially, crossed the inner membrane. Pulse-chase experiments showed that the MalE31 precursor was efficiently processed at the same rate as the wild-type protein and that no export interference or jamming was observed upon overexpression of *malE31*. Finally, transmission electron micrographs of cells overproducing MalE31 revealed the presence of small IBs between the inner and outer membranes (Figure 3B). Based on these data, it became clear that the mature MalE31 protein aggregates in the periplasm. As other aggregation-prone proteins, MalE31 could be purified from IBs, after a urea-solubilization step, and the renaturared protein *in vitro* exhibited complete maltose-binding activity. Thus, we concluded that the modified αβ turn, which is distant from the binding site, did not perturb the function but rather the periplasmic folding of MalE. Since these initial characterizations, several groups have used MalE31 as a protein model for cellular protein misfolding [71] or aggregation [89].

Compartmentalized cellular stress responses induced by malE31 *mutants*

To examine the alternative and competitive export and folding processes in the cytoplasm, we analyzed the cellular fates of MalE31 carrying either a wild-type or defective

signal sequences [90]. Since we never observed MalE31 misfolding in the cytoplasm when it was expressed with its wild-type signal sequence, we concluded that the MalE31 precursor enters rapidly into the Sec pathway. The mutational deletion or substitution in signal sequences has been shown to strongly decrease the efficiency of MalE export [91]. The presence of these altered signal sequences in combination with the defective folding *malE31* mutation resulted in the cytoplasmic accumulation of precursors. In the steady state, the cytoplasmic soluble fraction of MalE31, when expressed without a signal sequence (69%), is higher than the periplasmic soluble fraction of MalE31, when expressed with a wild-type signal sequence (5%). In the former case, the increased level of cytoplasmic chaperones (Dnak and GroEL), resulting from induction of the σ^{32} stress response, mainly suppresses the aggregation of MalE31 in the cytoplasm. Although overproduction of MalE31, carrying a wild-type signal sequence, induced an extracytoplasmic stress, the increased level of periplasmic Hsps failed to prevent the aggregation of MalE31. Thus, it appeared that there are insufficient or no general heat shock chaperones in the periplasm. Furthermore, we showed that aggregation of MalE31 in the periplasm did not induce a stress response via σ^{32} as it did when expressed with defective signal sequences, confirming that the stress responses induced by the presence of misfolded proteins in *E. coli* are compartmentalized and controlled by different pathways [90]. Technically, the determination of stress responses in *E. coli* is facilitated by using transcriptional gene fusions and allows the rapid discrimination of cellular location of the protein misfolding.

One physiological consequence for the cells overproducing MalE31 was to induce an extracytoplasmic stress response by increasing the expression of the heat shock protease DegP. Because the Cpx and σ^E regulatory systems both respond to the presence of misfolded exported proteins, we determined which pathway could be more specifically affected by MalE31. Therefore, by using two different transcriptional fusions which are independently regulated, either by the σ^E factor (P3*rpoH-lacZ*) or by CpxR (P*cpxP-lacZ*), we showed that overproduction of MalE31 induced the Cpx two-component system [49]. Because, the exact nature of inducers that activate this signaling pathway remained unclear, we examined whether the presence of periplasmic inclusion bodies is required. Interestingly, the Cpx system was also induced by the production of the destabilized MalE219 variant, which remains entirely soluble and functional in the periplasm (see below). This observation was further confirmed by producing truncated variants of MalE that did not accumulate at steady states [49]. It seems that the signal sensed by the Cpx pathway will be the presence of periplasmic unfolded or misfolded proteins rather than envelope alterations resulting from the formation of IBs.

Structure-dependence of the malE31 *mutation*

Despite a synergistic effect, the single $Ile^{33} \rightarrow$ Pro substitution was the major factor responsible for MalE31 aggregation. Thus, introduction of a proline residue in the first αβ turn of the N-domain could introduce either a conformational strain in the folded state or reduced conformational freedom of the polypeptide backbone in a critical folding intermediate state, because of the limited value of its dihedral angles. To probe genetically the structural role of this surface turn, we generated a library of random mutations at codons 32 and 33 of *malE*, and defective folding variants of MalE were screened [88]. Although many combinations of amino acid pairs replaced the native residues at both positions, underlining the passive role of this turn in determining the folded structure of MalE, the majority of critical folding substitutions, contained a proline residue at both positions. A second critical pair of residues involved the association of bulky hydrophobic residues (Phe, Leu, or Val) at position 32 and basic residues (Lys or Arg) at position 33. Close examination of the wild-type MalE crystal structure suggested that the combined negative effects of decreased conformational flexibility and exposed large hydrophobic side-chain residues at position 32, and a buried charged side-chain residue in the hydrophobic position 33, might create these defects. Since this study, the crystal structure of MalE31 was determined at 1.85Å resolution, and provided experimental evidence that MalE31 can attain a wild-type folded conformation *in vitro* and supports a model in which the structural effect of the *malE31* mutation is exerted at the level of folding intermediates, rather than of the native conformation [92].

The secondary and tertiary context-dependence of the *malE31* mutation was also assessed by probing the tolerance of an equivalent turn of the C-domain to the same double amino acid substitution. Among the four αβ turns of the MalE structure, we selected the turn connecting the helix αVII to strand βJ of the C-domain, and the corresponding $Gly^{220} \rightarrow$ Asp and $Glu^{221} \rightarrow$ Pro substitutions were introduced by site-directed mutagenesis [93], defining the MalE219 variant (Figure 3A). In contrast to MalE31, this modified-turn MalE219 variant was correctly folded and functionally active in the bacterial periplasm. Furthermore, *in vitro* unfolding/refolding experiments showed that both turn MalE variants were destabilized, but their intrinsic tendency to aggregate, correlated well with their periplasmic fates in *E. coli*, and arose from differences in refolding rates allowing long-term persisting folding intermediates of MalE31, probably kinetically trapped in an off-pathway folding reaction, to aggregate. These data supported the notion that the formation of this

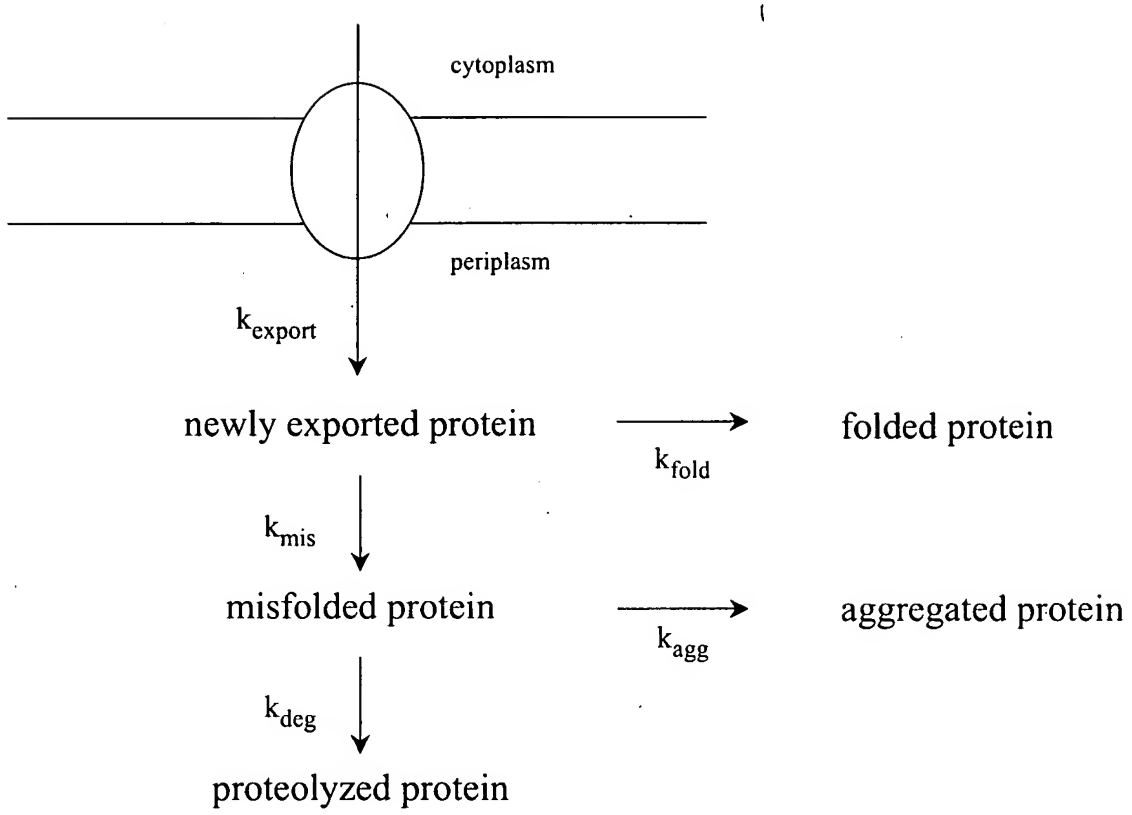


Figure 4
Kinetic competition between folding, degradation and aggregation in the periplasm. The model represents the export (zero order rate constant, k_{export}), folding (first order rate constant, k_{fold}), degradation (first order rate constant, k_{deg}), and aggregation reactions (second order rate constant, k_{agg}). In the postulated scheme, the newly exported protein is assumed to have two alternative pathways, either it correctly folds to give the folded protein or it proceeds via a side reaction leading to a misfolded protein. Once again, a kinetic competition is thought to occur between degradation and aggregation.

βαβ supersecondary structure of the N-domain is a rate-limiting step in the folding pathway of MalE. Indeed, previous intragenic suppressive mutations of export-defective signal peptides of MalE were identified in this structural element, and the analysis of refolding properties of the corresponding variants resulted in the same conclusion [94]. Together, these results argued that the information directing the polypeptide chain to the misfolding pathway is not only determined by a local amino acid sequence (the primary structure), but also by the tertiary structure surrounding the critical MalE31 substitution.

Cellular factors influencing inclusion body formation of MalE31

In a following series of experiments, we manipulated two cellular functions affecting protein folding pathways; the promoter activity of malE31 that modulates the rate of protein biosynthesis and the state of periplasmic DegP protease which determines cellular fates of misfolded MalE31 [95].

When expressed from its own promoter, the cellular level of MalE31 is controlled by the transcriptional activator MalT. Depending on the presence of either the wild type or constitutive MalT variant ($MalT^C$), the production of MalE31 varied from a very low level to a high and constitutive level in the host cells. We took advantage of

this regulation to modulate the biosynthetic rate of MalE31, and observed that at a low level of expression, MalE31 was rapidly degraded by the DegP protease, and that at a high level of expression, the misfolded protein formed IBs [95]. Surprisingly, the steaty-state level of soluble MalE31 was only dictated by the ratio of rate constants governing the misfolding and folding steps, and was independent on the promoter activity. The high production level of MalE31 in the *malT*^{C} context had two important consequences. First, increased amounts of aggregating MalE31 species increases the formation of IBs, the compact nature of which renders them resistant to proteases. Second, although overproduction of MalE31 increased the rate of DegP synthesis via the Cpx system [49], the majority of that increased amount of protease was localized in the insoluble fraction. Because DegP is intimately involved in the degradation of misfolded proteins in the periplasm, it was conceivable that DegP co-aggregated with misfolded MalE31 during the formation of IBs. All these data were quantified and analyzed by numerical simulations [95] on the working model presented in Figure 4. Because the kinetics of signal sequence processing of precursors were unaffected by MalE31 overproduction, export to the periplasm does not constitute a rate-limiting step for the periplasmic folding pathway of MalE31. The extent of aggregation is determined by both the rate of protein biosynthesis, which is related for MalE31 to the activity of its gene promoter, and protein folding or misfolding rates. At a high level of expression, second order aggregation dominates over first order folding. The fate of misfolded proteins depends obviously on the existence of cellular proteases. At the misfolding stage a second kinetic competition between degradation and aggregation might occur, and a further increase in the rate of expression (as in the *malT*^{C} background) will favor the aggregation reaction. At the time this model was proposed, we speculated that periplasmic chaperones and foldases might act directly on some kinetic partitioning steps.

FkpA suppresses the formation of MalE31 inclusion bodies

The potential role of the three known periplasmic PPIases was studied both *in vivo*, by overproducing them, and *in vitro* by recording the refolding kinetics of MalE31 in the presence of purified PPIases. Among these folding catalysts, only FkpA prevented the periplasmic aggregation of MalE31 [37]. Using active site and deletion variants, we demonstrated that the chaperone activity of FkpA was independent of its prolyl isomerase catalytic activity, but requires an intact dimerization N-domain [39]. *In vitro*, the presence of FkpA at stoichiometric amounts in the refolding buffer did not change the rate of the slow refolding phase of MalE31, but increased its relative amplitude as a chaperone does in general [37]. These results suggested that the ability of FkpA to increase the yield of soluble MalE31 can be explained by the blocking of the initial misfolding step through its binding to newly translocated proteins rather than to misfolded proteins which leads to the second competition between aggregation and degradation, as shown in Figure 4. Indeed, in the absence or in the presence of FkpA, the steady state level of MalE31 remained unchanged, and any modification in the balance between the relative rate of aggregation and degradation would lead to different levels, as we observed with the protease DegP. Thus, FkpA would bind newly exported MalE31 proteins in the periplasm, consistent with the definition of 'holder' chaperones, which can prevent the aggregation of unfolded protein without mediating protein reactivation.

Temperature effect on MalE31 inclusion body formation

Generally, the extent of aggregation is greater at higher temperature due to the strong temperature-dependence of the hydrophobic interactions, which dominate protein aggregation. Although MalE31 aggregated at 30°C, the formation of periplasmic IBs did not affect bacterial growth in rich medium [87]. In contrast, at 37°C, accumulation of IBs became very toxic [49]. Surprisingly, at 42°C, IBs did not accumulate and bacterial growth was restored. To explain this toxicity we hypothesized that some periplasmic or outer membrane proteins, which are essential for envelope biogenesis or cell division, could be incorporated into IBs and therefore lost to the cells. Indeed, DegP [95] and RseB [71] have previously been found trapped with MalE31 aggregates at 30°C. Under heat shock conditions no IBs were detected, thus we examined whether a temperature shift up could suppress the growth defect by growing cells at 37°C and then shifting them to 42°C. Interestingly, after transfer to 42°C the bacterial growth resumed, while the turbidity of cells left at 37°C decreased. We showed that the suppression of MalE31 toxicity after the temperature upshift resulted in the degradation of pre-existing IBs accumulated at 37°C. The only clearly documented heat shock protease which degrades misfoded periplasmic proteins is DegP. The synthesis of DegP at 42°C, as for other σ^{E} regulon members, is activated and its proteolytic activity also significantly increases with rising temperature [96]. These cumulative effects could contribute to a further increase of proteolytic activity in the periplasm at 42°C, and the degradation of MalE31 would be favored over aggregation in the model (Figure 4). In addition, because the steady-state level of wild-type MalE was slightly lower at 42°C [49], a further decrease in the rate of MalE31 synthesis, at this temperature, will also facilitate the degradation reaction. However, it has been reported that highly aggregated proteins were not proteolyzed by DegP *in vitro* [97], and we do not know if DegP is able to degrade IBs. Alternatively, we have no evidence that misfolded folding intermediates of MalE31 can dissociate from IBs and then be degraded by

DegP. Recent studies have revealed an unexpected dynamic process of IBs [98], from which aggregated proteins are steadily released or disaggregated to be further refolded or degraded by intracellular proteases [99]. Nevertheless, it is clear that under heat shock conditions, the progressive clearing of periplasmic IBs eliminated toxicity and restored bacterial growth.

Perspectives

Protein folding is one of the great unsolved problems of modern molecular biology. Not only is this question intrinsically fascinating, but it is closely tied to a wide range of applications and debilitating diseases, which are of critical importance in both biotechnology and molecular medicine. Most proteins that require chaperones to fold efficiently are poor folding models for biophysical experiments and, conversely, proteins suitable for *in vitro* folding studies do not require chaperone assistance. The complete characterization of a protein model, as MalE, and the availability of both *in vivo* and *in vitro* approaches, complementing each other, make it a particularly attractive system for mechanistic studies of this central issue. Our understanding of a number of cellular factors relevant to protein folding and assembly has increased enormously during the past five years, and quite remarkable progress has been made in the applications of this knowledge.

Although many alternative organisms and expression systems are now being developed, the most widely used host for high-level expression of recombinant proteins remains *E. coli*. Many important proteins for therapeutic or diagnostic applications including secreted enzymes contain disulfide bonds that are required for correct folding. Thus, studies on periplasmic protein folding have significant ramifications for biotechnology. However, many of these proteins are produced in biologically inactive and aggregated forms (IBs). Although solubilization and refolding procedures exist to obtain functionally soluble proteins, as well demonstrated with MalE31, they are lengthy processes, and not generalisable. For these proteins, modification of the amino acid sequences by mutagenesis, the folding code, may be one way to isolate a soluble form, and we can envision that genetic screens or selections for the rapid identification of correctly folded proteins expressed in the bacterial periplasm will be developed.

Other potential applications include the use of envelope folding factors as targets for the development of new antimicrobial drugs. Indeed, several proteins involved in periplasmic protein quality control are either essential or play an important role in the virulence of many pathogenic Gram-negative bacteria.

Acknowledgements

We are grateful to Emmett Johnson for critical reading of the manuscript. This work was supported by grants from the Institut Pasteur and the CNRS.

References

1. Goff SA, Goldberg AL: **Production of abnormal proteins in *E. coli* stimulates transcription of *lon* and other heat shock genes.** *Cell* 1985, **41**:587-595.
2. Wickner S, Maurizi MR, Gottesman S: **Posttranslational quality control: folding, refolding, and degrading proteins.** *Science* 1999, **286**:1888-1893.
3. Raivio TL, Silhavy TJ: **The sigmaE and Cpx regulatory pathways: overlapping but distinct envelope stress responses.** *Curr Opin Microbiol* 1999, **2**:159-65.
4. Pugsley AP: **The complete general secretory pathway in Gram-Negative bacteria.** *Microbiol Rev* 1993, **57**:50-108.
5. DeLisa MP, Tullman D, Georgiou G: **Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway.** *Proc Natl Acad Sci USA* 2003, **100**:6115-6120.
6. van den Berg B, Clemons WM Jr, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA: **X-ray structure of a protein-conducting channel.** *Nature* 2004, **427**:36-44.
7. Economou A, Wickner W: **SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion.** *Cell* 1994, **78**:835-843.
8. Luirink J, Dobberstein B: **Mammalian and *Escherichia coli* signal recognition particles.** *Mol Microbiol* 1994, **11**:9-13.
9. Danese PN, Silhavy TJ: **Targeting and assembly of periplasmic and outer-membrane proteins in Escherichia coli.** *Annu Rev Genet* 1998, **32**:59-94.
10. Sijbrandi R, Urbanus ML, ten Hagen-Jongman CM, Bernstein HD, Oudega B, Otto BR, Luirink J: **Signal recognition particule (SRP)-mediated targeting and Sec-dependent translocation of extracellular *Escherichia coli* protein.** *J Biol Chem* 2003, **278**:4654-4659.
11. Schierle CF, Berkmen M, Huber D, Kumamoto CA, Boyd D, Beckwith J: **The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway.** *J Bacteriol* 2003, **185**:5706-5713.
12. Valent QA, Kendall DA, High S, Kusters R, Oudega B, Luirink J: **Early events in preprotein recognition in *E. coli*: interaction of SRP and trigger factor with nascent polypeptides.** *EMBO J* 1995, **14**:5494-5505.
13. Hesterkamp T, Hauser S, Lutcke H, Bukau B: ***Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains.** *Proc Natl Acad Sci USA* 1996, **93**:4437-4441.
14. Eisner G, Koch H-G, Beck K, Brunner J, Müller M: **Ligand crowding at a nascent signal sequence.** *J Cell Biol* 2003, **163**:35-44.
15. Kumamoto CA: **SecB protein: a cytosolic export factor that associates with nascent exported proteins.** *J Bioenerg Biomembr* 1990, **22**:337-351.
16. Kusukawa N, Yura T, Uegushi C, Akiayama Y, Ito Y: **Effects of mutations in heat-shock genes *groES* and *groEL* on protein export in *Escherichia coli*.** *EMBO J* 1989, **8**:3517-3521.
17. Wild J, Altman E, Yura T, Gross CA: **DnaK and DnaJ heat-shock proteins participate in protein export in Escherichia coli.** *Gene Dev* 1992, **6**:1165-1172.
18. Bukau B: **Regulation of the *Escherichia coli* heat-shock response.** *Mol Microbiol* 1993, **9**:671-680.
19. Randall LL, Hardy SJ: **High selectivity with low specificity: how SecB has solved the paradox of chaperone binding.** *Trends Biochem Sci* 1995, **20**:65-9.
20. Paetzel M, Dalbey RE, Strynadka NCJ: **Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor.** *Nature* 1998, **396**:186-190.
21. Kadokura H, Katzen F, Beckwith J: **Protein disulfide bond formation in prokaryotes.** *Annu Rev Biochem* 2003, **72**:111-135.
22. Bader M, Muse W, Ballou DP, Gassner C, Bardwell JCA: **Oxidative protein folding is driven by the electron transport system.** *Cell* 1999, **98**:217-227.
23. Rietsch A, Belin D, Martin N, Beckwith J: **An *in vivo* pathway for disulfide bond isomerization in *Escherichia coli*.** *Proc Natl Acad Sci USA* 1996, **93**:13048-13053.

24. Kadokura H, Tian H, Zander T, Bardwell JCA, Beckwith J: **Snapshots of DsbA in action: detection of proteins in the process of oxidative folding.** *Science* 2004, **303**:534-537.
25. Liu J, Walsh CT: **Peptidyl-prolyl *cis-trans* isomerase from *Escherichia coli*: a periplasmic homolog of cyclophilin that is not inhibited by cyclosporin A.** *Proc Natl Acad Sci USA* 1990, **87**:4028-4032.
26. Horne SM, Young KD: ***Escherichia coli* and other species of the *Enterobacteriaceae* encode a protein similar to the Mip-like FK506-binding protein.** *Arch Microbiol* 1995, **163**:357-365.
27. Dartigalongue C, Raina S: **A new heat-shock gene, *ppiD*, encodes a peptidyl-prolyl isomerase required for folding of outer membrane proteins in *Escherichia coli*.** *EMBO J* 1998, **17**:3968-3980.
28. Rouviere PE, Gross CA: **SurA, a periplasmic protein with peptidyl-prolyl isomerase activity, participates in the assembly of outer membrane porins.** *Genes Dev* 1996, **10**:3170-3182.
29. Scholz C, Scherer G, Mayr LM, Schindler T, Fischer G, Schmid FX: **Prolyl isomerase do not catalyse isomerization of non-prolyl peptide bonds.** *Biol Chem* 1998, **379**:361-365.
30. Barnhart MM, Pinkner JS, Soto GE, Sauer FG, Langermann S, Waksman G, Frieden C, Hultgren SJ: **PapD-like chaperones provide the missing information for folding of pilin proteins.** *Proc Natl Acad Sci USA* 2000, **97**:7709-7714.
31. Missiakas D, Betton J-M, Raina S: **New components of protein folding in extracytoplasmic compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH.** *Mol Microbiol* 1996, **21**:871-884.
32. Schäfer U, Beck K, Müller M: **Skp, a molecular chaperone of gram-negative bacteria is required for the formation of soluble periplasmic intermediates of outer membrane proteins.** *J Biol Chem* 1999, **274**:24567-24574.
33. Chen J, Song J-L, Zhang S, Wang Y, Cui D-F, Wang CC: **Chaperone activity of DsbC.** *J Biol Chem* 1999, **274**:19601-19605.
34. Shao F, Bader MW, Jakob U, Bardwell JCA: **DsbG, a protein disulfide isomerase with chaperone activity.** *J Biol Chem* 2000, **275**:13349-13352.
35. Richarme G, Caldas TD: **Chaperone properties of the bacterial periplasmic substrate-binding proteins.** *J Biol Chem* 1997, **272**:15607-15612.
36. Matsuzaki M, Kiso Y, Yamamoto I, Satoh T: **Isolation of a periplasmic molecular chaperone-like protein of *Rhodobacter sphaeroides* f. sp. *denitrificans* that is homologous to the dipeptide transport protein DppA of *Escherichia coli*.** *J Bacteriol* 1998, **180**:2718-2722.
37. Arie J-P, Sassoon N, Betton J-M: **Chaperone function of FkpA, a heat shock prolyl isomerase, in the periplasm of *Escherichia coli*.** *Mol Microbiol* 2001, **39**:199-210.
38. Behrens S, Maier R, de Cock H, Schmid FX, Gross CA: **The SurA periplasmic PPIase lacking its parvulin domains functions *in vivo* and has chaperone activity.** *EMBO J* 2001, **20**:285-294.
39. Saul FA, Arié J-P, Vulliez-le Normand B, Kahn R, Betton J-M, Bentley GA: **Structural and functional studies of FkpA from Escherichia coli, a cis/trans peptidyl-prolyl isomerase with chaperone activity.** *J Mol Biol* 2004, **335**:595-608.
40. Bitto E, McKay DB: **Crystallographic structure of SurA, a molecular chaperone that facilitates folding of outer membrane porins.** *Structure* 2002, **10**:1489-1498.
41. Zarnt T, Tradler T, Stoller G, Scholz C, Schmid FX, Fischer G: **Modular structure of the trigger factor required for high activity in ptotein folding.** *J Mol Biol* 1997, **271**:827-837.
42. McCarthy AA, Haebel PW, Törrönen A, Rybin V, Baker EN, Metcalf P: **Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*.** *Nature Struct Biol* 2000, **7**:196-199.
43. Lipinska B, Zylicz M, Georgopoulos C: **The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase.** *J Bacteriol* 1990, **172**:1791-1797.
44. Strauch KL, Johnson K, Beckwith J: **Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature.** *J Bacteriol* 1989, **171**:2689-2696.
45. Erickson JW, Gross CA: **Identification of the σ^E subunit of *Escherichia coli* RNA polymerase: a second alternate σ factor involved in high-temperature gene expression.** *Genes Dev* 1989, **3**:1462-1471.
46. Clausen T, Southan C, Ehrmann M: **The HtrA family of proteases: implications for protein composition and cell fate.** *Mol Cell* 2002, **10**:443-455.
47. Krojer T, Garrido-Franco M, Huber H, Ehrmann M, Clausen T: **Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine.** *Nature* 2002, **416**:455-459.
48. Spiess C, Beil A, Ehrmann M: **A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein.** *Cell* 1999, **97**:339-347.
49. Hunke S, Betton J-M: **Temperature effect on inclusion body formation and stress response in the periplasm of *Escherichia coli*.** *Mol Microbiol* 2003, **50**:1579-1589.
50. Raivio TL, Silhavy TJ: **Periplasmic stress and ECF sigma factors.** *Annu Rev Microbiol* 2001, **55**:591-624.
51. Dong J, Iuchi S, Kwan H-S, Lu Z, Lin ECC: **The deduced amino-acid sequence of the cloned *cpxR* gene suggests the protein is the cognate regulator for the membrane sensor, CpxA, in a two-component signal transduction system of *Escherichia coli*.** *Gene* 1993, **136**:227-230.
52. Weber RF, Silverman PM: **The Cpx proteins of *Escherichia coli* K 12: structure of the CpxA polypeptide as an inner membrane component.** *J Mol Biol* 1988, **203**:467-478.
53. Nakayama S-I, Watanabe H: **Involvement of *cpxA*, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of *Shigella sonnei virF* gene.** *J Bacteriol* 1995, **177**:5062-5069.
54. Mileykosvskaya E, Dowhan W: **The Cpx two-component signal transduction pathway is activated in *Escherichia coli* mutant strains lacking phosphatidylethanolamine.** *J Bacteriol* 1997, **179**:1029-1034.
55. Snyder WB, Davis LJB, Danese PN, Cosma CL, Silhavy TJ: **Overproduction of NlpE, a new membrane lipoprotein, suppresses the toxicity of periplasmic LacZ by activation of the Cpx signal transduction pathway.** *J Bacteriol* 1995, **177**:4216-4223.
56. Jones CH, Danese PN, Pinkner JS, Silhavy TJ, Hultgren SJ: **The chaperone-assisted membrane release and folding pathway is sensed by two signal transduction systems.** *EMBO J* 1997, **16**:6394-6406.
57. Hung DL, Raivio TL, Jones CH, Silhavy TJ, Hultgren SJ: **Cpx signaling pathway monitors biogenesis and affects assembly and expression of P pili.** *EMBO J* 2001, **20**:1508-1518.
58. Raivio TL, Popkin DL, Silhavy TJ: **The Cpx envelope stress response is controlled by amplification and feedback inhibition.** *J Bacteriol* 1999, **181**:5263-72.
59. Raivio TL, Laird MW, Joly JC, Silhavy TJ: **Tethering of CpxP to the inner membrane prevents spheroplast induction of the Cpx envelope stress response.** *Mol Microbiol* 2000, **37**:1186-1197.
60. Danese PN, Snyder WB, Cosma CL, Davis LJB, Silhavy TJ: **The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP.** *Genes Dev* 1995, **9**:387-398.
61. Pogliano J, Lynch AS, Belin D, Lin ECC, Beckwith J: **Regulation of Escherichia coli cell envelope proteins involved in protein folding and degradation by the Cpx two-component system.** *Genes Dev* 1997, **11**:1169-1182.
62. Silverman PM: **Host cell-plasmid interactions in the expression of DNA donor activity by F^+ strains of *Escherichia coli* K-12.** *Bioessays* 1985, **2**:254-259.
63. Leclerc GJ, Tartera C, Metcalf ES: **Environmental regulation of Salmonella typhi invasion-defective mutants.** *Infect Immun* 1998, **66**:682-691.
64. Dorel C, Vidal O, Prigent-Combaret C, Vallet I, Lejeune P: **Involvement of the Cpx signal transduction pathway of *E. coli* in biofilm formation.** *FEMS Microbiol Lett* 1999, **178**:169-175.
65. Otto K, Silhavy TJ: **Surface sensing and adhesion of *Escherichia coli* controlled by the Cpx-signaling pathway.** *Proc Natl Acad Sci* 2002, **99**:2287-2292.
66. Raina S, Missiakas D, Georgopoulos C: **The *rpoE* gene encoding the σ^E (σ^{24}) heat shock sigma factor of *Escherichia coli*.** *EMBO J* 1995, **14**:1043-1055.
67. Rouvière PE, De Las Penas A, Mecsas J, Lu CZ, Rudd KE, Gross CA: ***rpoE*, the gene encoding the secon heat-shock sigma factor, σ^E, in *Escherichia coli*.** *EMBO J* 1995, **14**:1032-1042.
68. Mecsas J, Rouvière PE, Erickson JW, Donohue TJ, Gross CA: **The activity of σ^E, an *Escherichia coli* heat-inducible σ-factor, is**

modulated by expression of outer membrane proteins. *Genes Dev* 1993, **7**:2618-2628.
69. De Las Penas A, Connolly L, Gross CA: **The σ^E-mediated response to extracytoplasmic stress in *Escherichia coli* is transduced by RseA and RseB, two negative regulators of σ^E.** *Mol Microbiol* 1997, **24**:373-385.
70. Missiakas D, Mayer MP, Lemaire M, Georgopoulos C, Raina S: **Modulation of the *Escherichia coli* σ^E(RpoE) heat-shock transcription-factor activity by the RseA, RseB and RseC proteins.** *Mol Microbiol* 1997, **24**:355-371.
71. Collinet B, Yuzawa H, Chen T, Herrera C, Missiakas D: **RseB binding to the periplasmic domain of RseA modulates the RseA: sigmaE interaction in the cytoplasm and the availability of sigmaE.RNA polymerase.** *J Biol Chem* 2000, **275**:33898-33904.
72. Ades SE, Connolly LE, Alba BM, Gross CA: **The Escherichia coli σ^E-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-sigma factor.** *Gene Dev* 1999, **13**:2449-2461.
73. Alba BM, Leeds JA, Onufryk C, Lu CZ, Gross CA: **DegS and YaeI participate sequentially in the cleavage of RseA to activate the σ^E-dependent extracytoplasmic stress response.** *Gene Dev* 2002, **16**:2156-2168.
74. Walsh NP, Alba BM, Bose B, Gross CA, Sauer RT: **OMP peptide signals initiate the envelope-stress response by activating DegS protease via relief of inhibition mediated by its PDZ domain.** *Cell* 2003, **113**:61-71.
75. Dartigalongue C, Missiakas D, Raina S: **Characterization of the *Escherichia coli* sigma E regulon.** *J Biol Chem* 2001, **276**:20866-20875.
76. Dartigalongue C, Loferer H, Raina S: **EcfE, a new essential inner membrane protease: its role in the regulation of heat shock response in *Escherichia coli*.** *EMBO J* 2001, **20**:5908-5918.
77. Danese PN, Silhavy TJ: **The σ^E and the Cpx signal transduction systems control the synthesis of periplasmic protein-folding enzymes in *Escherichia coli*.** *Genes Dev* 1997, **11**:1183-1193.
78. Boos W, Shuman H: **Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation.** *Microbiol Mol Biol Rev* 1998, **62**:204-229.
79. Shuman HA: **Active transport of maltose in *Escherichia coli* K12.** *J Biol Chem* 1982, **257**:5455-5461.
80. Schatz PJ, Beckwith J: **Genetic analysis of protein export in *Escherichia coli*.** *Ann Rev Genet* 1990, **24**:215-248.
81. Bassford PJ: **Export of the periplasmic maltose-binding protein of *Escherichia coli*.** *J Bioenerg Biomembr* 1990, **22**:401-439.
82. Randall LL, Hardy SJ: **Correlation of competence for export with lack of tertiary structure of the mature species: a study in vivo of maltose-binding protein in *E. coli*.** *Cell* 1986, **46**:921-8.
83. Prinz WA, Spiess C, Ehrmann M, Schierle CF, Beckwith J: **Targeting of signal sequenceless proteins for export in *Escherichia coli* with altered protein translocase.** *EMBO J* 1996, **15**:5209-5217.
84. Spurlino JC, Lu GY, Quiocho FA: **The 2.3-Å resolution structure of the maltose-or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis.** *J Biol Chem* 1991, **266**:5202-5219.
85. Duplay P, Szmelcman S, Bedouelle H, Hofnung M: **Silent and functional changes in the periplasmic maltose-binding protein of *Escherichia coli* K12.** *J Mol Biol* 1987, **194**:663-673.
86. Duplay P, Hofnung M: **Two regions of mature periplasmic maltose-binding protein of *Escherichia coli* involved in secretion.** *J Bacteriol* 1988, **170**:4445-4450.
87. Betton J-M, Hofnung M: **Folding of a mutant maltose binding protein of *E. coli* which forms inclusion bodies.** *J Biol Chem* 1996, **271**:8046-8052.
88. Betton J-M, Boscus D, Missiakas D, Raina S, Hofnung M: **Probing the structural role of an $\alpha\beta$ loop of maltose-binding protein by mutagenesis: heat-shock induction by loop variants of the maltose-binding protein that form periplasmic inclusion bodies.** *J Mol Biol* 1996, **262**:140-150.
89. Wigley WC, Stidham RD, Smith NM, Hunt JF, Thomas PJ: **Protein solubility and folding monitored *in vivo* by structural complementation of a genetic marker protein.** *Nat Biotechnol* 2001, **19**:131-136.
90. Betton J-M, Phichith D, Hunke S: **Folding and aggregation of export-defective mutants of the maltose-binding protein.** *Res Microbiol* 2002, **153**:399-404.
91. Bedouelle H, Bassford PJ, Fowler AV, Zabin I, Beckwith J, Hofnung M: **Mutations which alter the function of the signal sequence of the maltose binding protein of *Escherichia coli*.** *Nature* 1980, **285**:78-81.
92. Saul FA, Mourez M, Vulliez-Le Normand B, Sassoon N, Bentley GA, Betton J-M: **Crystal structure of a defective folding protein.** *Protein Sci* 2003, **12**:577-85.
93. Raffy S, Sassoon N, Hofnung H, Betton J-M: **Tertiary structure-dependence of misfolding substitutions in loops of the maltose-binding protein.** *Protein Sci* 1998, **7**:2136-2142.
94. Chun S-Y, Strobel S, Bassford PJ, Randall LL: **Folding of maltose-binding protein.** *J Biol Chem* 1993, **268**:20855-20862.
95. Betton J-M, Sassoon N, Hofnung M, Laurent M: **Degradation versus aggregation of misfolded maltose-binding protein in the periplasm of *Escherichia coli*.** *J Biol Chem* 1998, **273**:8897-8902.
96. Skorko-Glonek J, Krzewski K, Lipinska B, Bertoli E, Tanfani F: **Comparaison of the structure of wild-type HtrA heat shock protease and mutant HtrA proteins.** *J Biol Chem* 1995, **270**:11140-11146.
97. Kim KI, Park S-C, Kang SH, Cheong G-W, Chung CH: **Selective degradation of unfolded proteins by the self-compartmentalizing HtrA protease, a periplasmic heat shock protein in *Escherichia coli*.** *J Mol Biol* 1999, **294**:1363-1374.
98. Carrio MM, Villaverde A: **Construction and deconstruction of bacterial inclusion bodies.** *J Biotechnol* 2002, **96**:3-12.
99. Carrio MM, Corchero JL, Villaverde A: **Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms.** *Biochim Biophys Acta* 1999, **1434**:170-6.